CORRELATES OF EARLY CHILDHOOD HEALTH IN ETHIOPIA: *APPLICATION OF STRUCTURAL EQUATION MODELING WITH LATENT VARIABLES*

BY:

Biniyam Mekasha Yemane

A thesis Submitted to the School of Graduate Studies of Addis Ababa University in Partial Fulfillment of the Requirement for the Degree of Masters of Science in Economics (Economic Policy Analysis)

July 2009

Addis Ababa

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

"Correlates of Early Childhood Health in Ethiopia: Application of Structural Equation Modeling with Latent Variables."

By

Binyam Mekasha Yemane

Approved by the Board of Examiners:

Tassew Woldehanna	
Advisor	**Signature**
Hussien Hamda	
Examiner	**Signature**
Getachew Yoseph	
Examiner	**Signature**

Table of Contents

Pages

List of Tables

Pages

List of Figures

Pages

List of Appendices

Acronyms

ADF	Asymptotic Distribution Free estimation
AGFI	Adjusted Goodness of Fit Index
AMOS	Analysis of MOments Structures
CSA	Central Statistics Authority
EDRI	Ethiopian Development Research Institute
FA	Factor Analysis
GDP	Gross Domestic Product
HDI	Human Development Index
LISREL	Linear Structural Relations
LR	Likelihood Ratio
MDG	Millennium Development Goal
MIMIC	Multiple Indicators and Multiple Causes
ML	Maximum Likelihood
MOFED	Ministry of Finance and Economic Development
NCHS	National Center for Health Statistics
OAU	Organizations of African Unity
PCA	Principal Component Analysis
PLS	Partial Least Square
PRELIS	PRE-processor for LISREL
RMR	Root Mean squared Residual
RMSEA	Root Mean Squared Error Approximation

SEM	Structural Equation Modeling
SIMPLES	Simple LISREL
SSA	Sub-Saharan Africa
UK	United Kingdom
UN	United Nations
UNDP	United Nations Development Programme
UNICEF	United Nations Children's Fund
WLS	Weighted Least Square
WMS	Welfare Monitoring Survey
YLS	Young Lives Study

Abstract

Child health status is an important determinant of adult productivities and earnings; thus it represents an important channel of intergenerational socioeconomic mobility. It is also recognized that health is a variable inherently unobservable and if observed has many measurement errors. This thesis used the special case of structural equation models with latent variables known as the Multiple Indicators and Multiple Causes (MIMIC) to investigate the determinants of child health and avoid the problem of health unobservability by using three indicators. The source of the data is the Young Lives Project's household survey. The novelty of the empirical analysis is the simultaneous estimation of equations for both permanent and transitory health using cross-sectional designs. The results of this paper show that parental education, household size, safe drinking water source of the household, household overall economic status and vaccination availability are among the important determinants of both permanent and transitory child health. The study also showed that there is strong correlation between transitory and permanent child health in that a good permanent health state determines the transitory health state significantly. The study finally concluded that any policy measure towards improving the health status of children has to take in to account the above mentioned factors as target variables.

Keywords: Correlates, Child Health, Ethiopia, Transitory, Permanent, Structural Equation Modeling (SEM), MIMIC Model

Chapter I

1. Introduction

1.1 Introduction and Background

Childhood is a critical period of development in life. Poor health during this period may cause serious damages all throughout the life cycle. It is observed that usually poor children grow up to be a poor adult in that they pass on their poverty to their children. This intergenerational transfer of economic status starts as early as childhood (Case *et al.,* 2002; Currie and Moreti, 2003; Currie and Stablie, 2003), or in the uterus (Almond, 2005; Case *et al.*, 2005). To break this vicious cycle, children must be provided with the appropriate food, shelter, health care, education, other public services and they must have a voice in the community (UNICEF, 2005).

According to Behrman and Deolalikar (1988) health and nutrition may be channels through which productivity and distributional goals of developing societies may be pursued effectively if, as often hypothesized, the productivity of low income persons in work and human capital formation is positively affected by health and nutrition. Thus, a sensible developmental policy of any country must obviously aim at providing for a level of calorie intake that will permit the full productivity and work output from its labor force, and a level of growth and development for its children that represents the fullest expression of their genetic potential.

Child health has become a key indicator of economic development. Among the eight Millenium Development Goals (MDGs) that were adopted by the 189 members of the United Nations (UN) in 2000, at least four are directly related to child health or nutritional status

(Todaro and Smith, 2005).[1] In addition to being a development indicator itself, child health is also closely associated with other development indicators, such as adult health, educational attainment, productivity, and income (Currie and Hyson, 1999; Currie and Madrian, 1999; Persico *et al.*, 2004; Case *et al*,, 2002, 2005; Behrman and Rosenzweing, 2004). Despite its importance, however, little is known about the causes of good or poor child health (Yuyu and Hongbin, 2009).

However, in most cases children are neglected in development planning or seen as special cases and thereby 'tagged' on the development agenda (Harper *et al*, 2003). As diversity exists at all levels of society, policy processes have to be adopted to tackle differences, children whilst diverse with in the group, cannot be labeled as special group, not least because they comprise almost half of the world's population and their interests are inextricably linked with those of the adults (Harper and Marcus, 1999/00). This has been recognized to some extent in policy where the welfare of children is highly linked to the welfare of adults and the communities in which they live.

The health status of children has been given greater concern recently-for poor health in childhood has so many implications. For example, temporary shortfalls in health or nutrition can have lasting and irreversible effects when they occur during childhood, a period of significant development (Jensen and Richter, 2001). It is argued that health investment may be one channel through which poverty and disadvantage are transmitted across generations, as children of poor parents receive less health investments, which in turn may reduce future earning capacity.

[1] These four goals are to reduce child mortality, improve the mother's health, combat diseases such as AIDS and Malaria, and eradicate hunger.

Other studies (Case *et al*, 2002) indicated that in addition to the direct welfare and financial costs of illness, poor childhood health results in lower levels of human capital accumulation. Because, less healthy children spend more days in bed and miss more school they may have no further opportunities to learn key skills and knowledge for the future (Glewwe, Jacoby and king, 2001).

In general, results from the literature show the importance of adequate health for children and therefore the determinants behind these outcomes must be understood.

Most of the measurements used to capture this health status of children, however, are problematic for two reasons at least at this stage. First, it is multidimensional; second, measurement error in health is likely to be related to income and labor market outcomes (Strauss and Thomas, 1998). As a result, anthropometric measurements are used as general indicators of short term and long term health outcomes in most of the literature (see for example Behrman and Deolalikar, 1988; Senaur and Gharcia, 1991; Strauss and Thomas, 1995; Glewwe, 1999; Schultz, 2002; Tirfe, 2006; Ayalew, 2006)

This paper hence explored factors determining child health in Ethiopia and estimated a special case of structural equation models known as Multiple Indicators-Multiple Causes (MIMIC) models. The model is estimated with a computer package known as LISREL (Linear Structural Relations). Using the same methodology Wolfe and Behrman (1984), Joreskog and Sorbom (1989), Kiiskinen (2003), Shehzad (2004, 2006)), and showed how the issue of health unobservability can best be captured by employing MIMIC models in which unobserved (latent) variables are caused by several observed factors and indicated by several observed indicators.

1.2 Statement of the problem

It is now widely acknowledged that higher investment in human capital such as health of children may not only improve their current welfare but also enhance their opportunities for escaping from poverty as adults. Improved health improves the mental and physical capabilities of children and thereby their future earning capacity. In a wider context this may mean contributing to the potentials of the economic growth and poverty reduction efforts of a given country like Ethiopia.

More generally, improving poor health status and hence reducing child poverty is a key goal of most developing countries. A number of studies have documented the wide range of adverse economic and social consequences of child poverty. For instance, malnutrition during infancy and childhood substantially raises vulnerability to infection and disease and increases the risk of premature death. It is also believed to impair cognitive achievement, labor productivity during adulthood, and lifetime earnings (Glewwe, Jacoby, and King, 2001; Glewwe and King, 2001). Thus, combating child malnutrition and hence poor early childhood health is of central importance to the economic and social welfare of countries.

For countries to engage in activities that improve child health with the right set of interventions, their policy makers need to have a better understanding of its economic, social, and policy correlates. While there have been several studies that have analyzed the socioeconomic correlates of child health as indicated by various indicators such as infant mortality and malnutrition, they suffer from two major limitations.

First, they do not focus enough on direct and indirect policy interventions, such as improved infrastructure, that could have as large effects on childhood health as direct individual, household and community characteristics. Second and most importantly, previous studies have almost exclusively concerned themselves with estimating the mean effect of variables such as a child's sex, the schooling of its mother, and household income. Such estimates miss a point that is crucial for policy makers – viz., that socioeconomic background variables and policy interventions may affect child health. In an effort to identify the correlates of early childhood health, this study has given due emphasis to the identification of policy variables that significantly affect child health.

Moreover, theoretical and empirical work relating to child health identifies various problems in health measurement that are not taken in to account in previous studies, these broadly include:

(i) unobservability of health status

(ii) problem of differentiating between different health states and

(iii) if health is measured through various indicators then the problem of measuring these inherently imperfect health indicators.

Behrman and Deolalikar (1988) examine that "*problems relating to health measurement are manifold and often result in certain controversies and ambiguities in the literature*". Grossman (1972) presents a model of health determination in which health is seen as a durable capital stock that depreciates overtime. However, Grossman also shows that "*the stock of health... is a theoretical concept, one that is difficult to quantify empirically*".

Studies by Behrman and Wolfe (1987), Van Doorslaer (1987), and Wagstaff (1993) identify problems of unobservability and suggest ways to solve them. The studies suggest using latent variables to overcome unobservability and employ micro-representations of health status. Behrman and Deolalikar (1988) argue that due to the imperfect nature of indicators, a number of indicators should be used to illustrate health. Hence, this paper is believed to fill this gap in the Ethiopian case. Van Doorslaer (1987) tried to distinguish between Grossman's stocks and flow concepts of health at the empirical level. However, his results relating to transitory health were peculiar: the coefficients of determinants were mostly of the opposite signs casting doubt on the relationship between different health states.

Most empirical research related to health measurement has therefore, been constrained by the issue of unobservability of health status and appropriate estimation techniques. This study addressed the complex issue of child health measurement and used latent variables to overcome health unobseravbility and hence fills the gap. To take care of the imperfect nature of health indicators, a standardized reference has been used for multiple health indicators.

The following specific issues have been addressed: *Child health is unobservable but can be represented by several observed health indicators* and *different observed health indicators represent different underlying health states.*

In this paper, a structural equation model of determinants of child health will be developed by using the Young Lives one year old children since this is the period when children are most vulnerable to disease and malnutrition. The approach of this study in dealing explicitly with the determinants of child health status is in the spirit of van de Ven and van der Gaag (1982); Wolfe and Behrman (1984); Joreskog and Sorbom (1989); Kiiskinen (2003); Shehzad (2004, 2006); and Giuffrida *et al* (2005). We treat transitory and permanent child health status as unobservable latent variables that are determined simultaneously. Identification of the socioeconomic and policy determinants of child health has been made possible by estimating the special case of structural equation modeling, i.e. Multiple Indicators and Multiple Causes (MIMIC) model.

The specification of the model followed the static theoretical framework. The use of cross-sectional analysis does not utilize the full capacity of data at hand, but is seen as an important first step in this type of analysis for two reasons:

1. Most similar studies published earlier have used cross-sectional and hence it allows a broader comparison of the results.

2. Including the determinants of both transitory and permanent health in the system makes it more complex than any previous studies. Coupled with multiple measurements for some key variables it is feasible to test the model before adding complexity by introducing dynamic aspects to the analysis.

1.3 Objectives of the Study

In general, the objective of this research is identifying the important factors behind poor early childhood health in Ethiopia with due emphasis to identifying the policy variables by introducing a new type of measuring health with multiple indicators.

Specifically, the objectives of the paper are;

(i) Identifying the correlates of early childhood health by making distinction between transitory and permanent health following Grossman's (1972) concept of health *stock* and *flow*.

(ii) Identifying link between the children's temporary and permanent health status;

(iii) Identifying policy interventions effective in reducing poor child health status.

In order to achieve the set objectives, the study employed Structural Equation Models with latent variable whose special case of Multiple Indicator Multiple Causes (MIMIC) with single and two latent variables is applied.

1.4 Significance of the study

Most of the studies in the area of child health in Ethiopia lack agreement about the relative importance of factors. However, in addition to this disagreement of results, the problem of early childhood poor health has not been well represented/measured and as a result the identification of factors was extremely dependent up on the indicators used in the particular study. This fact has been confirmed by theoretical and empirical work relating to child health that identified various problems in health measurement that are not taken it to account in previous studies including most importantly unobservability of health status.

To my knowledge, the methodology used in this paper is less utilized in the health literature especially in the African context. Hence, this paper is believed to fill this gap. From the structural equation modeling this paper applied the special case of Multiple Indicators and Multiple Cause (MIMIC) and estimated it with LISREL[2] software package.

However, LISREL modeling is still believed to be very complex and not quite often taken up by researchers. This study is, therefore, a break through to this complex type of analysis and is believed to be the stepping stone to the broader aspect of health through this modeling strategy that has been successfully breaking the backbone of most problems in economic measurement.

1.5 Limitation of the Study

This research is entirely dependent upon quantitative data. As a result it was difficult to give more sensible explanations for some less-intuitive results that came up, and this limited the analysis which could have been possible had there been qualitative information supplementing the research.

Moreover, although the selected MIMIC models are congruent with the data, and plausible, they remain one set of many potential plausible MIMIC models that have not been tested. Nonetheless, the selected models are informative about the determinants of child health. In summary, models could be improved if data were collected specifically to conduct MIMIC models, and the information provided would be enhanced if further waves of data were applied.

[2] See appendix B for more about the software.

1.6 Organization of the study

The paper is organized as follows. The thesis consists of five main chapters. Chapter 2 provides a comprehensive review of the existing literature on health and child health models and empirical studies of determinants of child health in Ethiopia and the rest of the world. The review is used to set out three "study questions" for the thesis to address. These relate to the determinants that are in play in influencing child health; the relationship between temporary and permanent child health, and the effective policy tools in reducing poor health state of children. Chapter 3 briefly specifies the methodology and the model which the later empirical analysis will be based upon. Chapter 4 presents empirical analysis of cross-sectional model. The summary, conclusions and policy implications are summarized in Chapter 5.

CHAPTER II

Review of Related Literature

2.1 Conceptual Framework and Measurement: in light of child health

In almost all of the literature related to child health the measurements used are directly related or same as the measurements used for malnutrition, that are the anthropometric measures. The word 'anthropometry' is generally meant to represent the measure of people's growth indicators such as weights and heights (related to their age and sex). It is used for growth assessment and is believed to be a single best measurement that defines the health/nutritional status of a child (Ayalew, 2006). However, the concept of health is multidimensional and all-around such that it is an outcome of complex interactions of social, economic, cultural and political factors. The multidimensionality has led to the difficulty of providing a commonly acceptable and unambiguous measure of health.

According to Blackman and Litchfield (2001), child health goes to cover issues of child poverty. Child health should therefore be measured by making use of indicators on different dimensions of welfare. Being in an ill state of health cannot explain everything. Increased morbidity and mortality from illness; inadequate housing and homelessness; unsafe environment; social discrimination and exclusion; and lack of participation in decision making in civil, social and cultural life have been taken as among the various manifestations of lower health status.

2.2 Theoretical Literature

It is essential to have a theoretical framework for the determinants of health in order to analyze the variables in an organized manner and to be able to interpret empirical studies (Behrman and Deolalikar (1988). In other words, for examining the determinants of child health it is important to see how decisions of households affect child health for given household assets and community endowments in a theoretical framework (Shehzad, 2004; 2006). From the Grossman's (1972) model with modification to children context, parents derive utility from good child health and ill-health. In this framework, utility from health is maximized under two set of constraints that are

(i) Resource constraint and

(ii) Production function constraint.

The production function for health is a technical relationship and shows that health output is produced with the help of certain inputs.

According to Grossman (1972), the argument of the gross investment production function includes medical care, time input and education. However, Behrman and Deolalikar (1988) examine that it is important to realize that our knowledge of the technical relations determining health is quite primitive. This is because health output may be greater if productivity is greater in a production function context.

The literature on health and child health reveals that three types of relationships can be identified as stated in Shehzad (2004, 2006): demand function, health production function and hybrid functions. The analysis of Behrman and Deolalikar (1988) examined however, that the reduced form demand relations do not provide much information about the structural coefficients (e.g education, time inputs, birth control effects, etc.), but provide a consistent framework within which to examine the impact of changes in market prices and endowments on health.

In this particular paper an estimate was provided with a mixed demand-production function for child health in Ethiopia. This is because child health is assumed to be determined by some inputs from the structural production function (such as family size and medical care) and some variables from the reduced form demand relations such as income as represented by the wealth index and household quality as indicated by the household quality index.

The United Nations Children's Fund's framework for the causes of child malnutrition (UNICEF, 1998) and the subsequent extended model of care can also be used contextualizing them to health as a conceptual framework guiding the empirical analysis. This is because it presents a useful generalized understanding of how child health are the outcomes of a multi-sectoral development problem that can be most effectively analyzed in terms of immediate, underlying and basic causes. The framework also showed that these determinants are determined by other basic determinants such as available resource of a country which are in turn affected by political and social factors (Smith and Haddad, 2000).

2.3 Empirical Literature

2.3.1 Determinants of Child Health: *a review of studies from the rest of the world*

Health in general is of interest to economists, first because it is an important element of wellbeing, and second because it is a component of human capital, and as such is of major importance for growth and development. In poor countries, where physical jobs tend to be more abundant, health may be more important than education in determining labor productivity (Kimhi, 2002).

The literature on health and economic development has been surveyed by Behrman and Deolalikar (1988), and more recently by Strauss and Thomas (1998). There is a sizeable body of empirical evidence showing that health is determined by family decisions. Studies have also shown that health is directly affected by nutritional intakes, while nutritional intakes are clearly affected by factors both outside and within the family. Jensen and Richter (2001) show that infant health is determined by maternal behavior. This behavior, however, seems to respond in turn to unanticipated health outcomes among children as well as to persistent health factors. Strauss (1990), and Thomas and Strauss (1992) show that health depends on the source of drinking water.

Most of these studies use anthropometric measures as proxies for health status. These are objective indicators, based on actual measurement of certain body properties such as height and weight, as briefly discussed in section 2.1. In developing countries, these reflect the interaction of nutrition and environmental variables such as infection. Several studies use self-reported indicators of health status, such as days ill and the ability to perform Activities of Daily Living (ADL). While day's ill may suffer from reporting errors, ADL-based measures are known to be valid and reliable. Wolfe and Behrman (1984) using similar methodology as this paper treated health status as a latent variable and estimated a system of equations including health status, household wealth and health care utilization. Alderman and Garcia (1994) show that health indicators affect nutritional status, while other studies have shown that nutritional status affects days ill. Both use instrumental variables because both health and nutritional status are determined endogenously.

Most studies reported that child characteristics such as age, sex and birth order are important determinants. Some of them found that child height for age and weight for age vary substantially with age whereby poor health/malnutrition rises with age in the first two years and levels off later (See Strauss, 1990; Glewwe, 1999; Ayalew, 2006). The above results did not find strong evidence of gender discriminations against females. In fact Sahn and Stifel (2002), Barrera (1990), Senaur and Garcia (1991) found girls achieving better outcomes than boys.

Parental and household characteristics such as parental age, education level, health knowledge, height, household size, household resources, etc are all important in child health outcomes. Specifically, most of the studies reported maternal education to have a positive and significant effect on child health/nutrition (See Barrera, 1990; Thomas *et al.*, 1990a; Senaur and Garcia, 1991; Escobal, 2005).

However, there is a controversy over this result in that Straus (1990) found that the effect of both mother's education and father's education are non linear and the impact was found negligible but maternal education was found to have a greater effect than fathers. Alderman

(1990) on the other hand, showed that education of the father has negative influence while the positive influence was not significant suggesting the low quality of education for women.

Studies such as Barrera (1990), Thomas *et al.* (1990b) Escobal *et al.* (2005) tried to see the role of maternal schooling and its interaction with public health programs in child health production and found significant relationship.

Household wealth such as income and household durables are believed to be one of the important factors in the production of child health. Nevertheless there is lack of consistency across studies over the significance of these variables. For instance, Thomas *et al.* (1990b) found total income to have a positive and significant effect on child height. However, on the contrary, as food availability is being one of household resources, both Alderman (1990) and Maxwell *et al.* (2000) did not find it to be a significant determinant rather they found female education as a substitute for income.

Since the number of household members has a serious implication in both the household resource share and in creating congestion effects on the child health status, it is a very important determinant. Senaur and Garcia (1991) found it to have a significant positive impact on height of children. The authors argued that this could be because household full income is a function of wage rates and the number of economically active family members. On the contrary, higher household size could also affect the health status of children by creating congestion effect and easy transmission of communicable disease throughout the family and hence lower child health status (see Alderman, 1990 and Shehzad, 2006).

With regards to the impact of community characteristics or environment, Attanasio *et al* (2005) used an estimation of regression using OLS and 2SLS, and found that having a hospital in the community improves children height. On the other hand, similar to that of Escobal *et al* (2005) and Barrera (1990) the coverage of the piped water network positively influences child health if the parents have some education.

2.3.2 Determinants of Child Health: *in Ethiopia*

The prevalence of poor child health in Ethiopia has been and is still very high for many years. As part of Welfare Monitoring Survey, the Central Statistics Authority of Ethiopia is periodically providing data on health status of children as indicated by nutritional status every two years since 1996. According to the Welfare Monitoring Survey of 2004 report, about 10.4 million children between the ages of 3-59 months are considered for anthropometric measurements. The results of successive surveys indicated that over time there is a tremendous decrease in the rate of malnutrition measured by percentage of stunting and underweight throughout the country.

In Ethiopia various studies have shown that the child characteristics including the age, gender and birth order are very important determinants of child health in some as indicated by child malnutrition (See Alemu *et al.*, 2005a, Tekie *et al.*, 2003 Tassew *et al.*, 2008). The Studies by Woldemariam and Timotious (2002); Alemu *et al.* (2005a) concluded that the probability of a child being underweight or stunted increases significantly as the child gets older. However, Christiaensen and Alderman's (2001) study found that a child's standardized height deteriorates up to the age of three, and slightly improves afterwards. Alemu *et al.* (2005b) using the Young Lives sample of eight year old children found that children have higher weight-for-height with the rate of increase decreasing as children grow older.

Studies by Bilisuma (2004), Ayalew (2006), Christiaensen and Alderman's (2001) and Alemu *et al.* (2005b) among others have shown that boys are indicated to be more vulnerable to ill health/malnutrition than girls. Among other justifications, Alemu *et al.* (2005a) argue that it could be due to genetic differences between male and female children and due to girl's greater access to food through their gender-ascribed role in contributing to food preparation. As opposed to this Silva (2005) did not find the coefficient on child's gender to be significant suggesting the absence of gender bias.

Alemu *et al.* (2005a) using the Young Lives children did not find birth order to be a significant determinant of malnutrition, a result at odds with that of Woldemariam and Timotios (2002).

In a typical developing country like Ethiopia, parental education is expected to improve the health status of children for reasons that higher level of education is associated with higher job opportunity and thereby higher income availability to the household and greater utilization of available health care services and better health promoting behaviors. However, findings show that the impact and direction of parental schooling on the determination of health status of children in Ethiopia is inconclusive.

Female's education has been found to be significantly associated with better health status in some studies such as Woldemariam and Timotios (2002), Alemu et al. (2005b) and Silva (2005) while both mother's education and father's education were found to be significantly associated with better nutrition status in Christiaensen and Alderman's (2001) study conducted using the three consecutive welfare monitoring surveys. As opposed to this a study by Alemu et al., (2005a) reported that there is a negative relationship between the education level of parents and child health.

Bilisuma (2004) estimated the determinants of child stunting in Ethiopia using both OLS and probit models. The probit result shows that primary and university level education of head of the household do have a significant positive impact on the probability of children being stunted. She also showed that having a secondary level of education decreases the chance of a child being stunted and it is significant.

In the Ethiopian child health literature also there is a contradictory result on the impact of household size on child health. Christiansen and Alderman (2001) found a result that reported a positive relationship while Sentayehu (2004) reported otherwise.

Most of the studies in Ethiopia found the significance of the importance of household wealth in determining the child health/nutrition (Alemu *et al.*, 2005b; Bilisuma, 2004; Christiansen and Alderman, 2001). On the other hand, children in those households that had a food shortage are found to be affected (Dercon and Hoddinot, 2003; Alemu *et al.* (2005a) though the direction and magnitude differs across different age and gender groups.

It is expected that the community characteristics in which the household is living in matters significantly in determining the health status of children. Consistent with expectations Silva (2005) found that household's own access to water and sanitation have negative coefficients. However, Silva (2005) emphasized that the external impact of access to water and sanitation facilities diminishes as the proportion of households in the community with access to water and sanitation increases. Possessions of a tap and a flush toilet have been found to have a positive effect on child height in Christiansen and Alderman (2001). However, access to other sources of drinking water was not found to positively affect children's height (see also Woldemariam and Timotios, 2002, and Alemu *et al.*, 2005a).

2.4 Latent Variables

It is impossible to date the first use of latent variables (Bollen, 2002). The idea that observable phenomena are influenced by underlying and unobserved cause is at least as old as religion, where unseen forces affect real world events. In the more secular sphere of everyday living, latent variables find wide applications. From the response to "how are you feeling today?" to the description of a worker as "efficient" or a student as "bright" such abstract elude direct measurement. These examples illustrate the common practice among humans to explain, to understand, and to sometimes predict events based on the role of concepts that are not directly observable.

Thus many of the phenomenon that we wish to explain are either measured with error or not directly measurable in economics and psychology (Weeks, 2001). Other examples include intelligence, education, health and institutional change. Perhaps the most prevalent example in economics is based upon Friedman's (1957) (as cited in Weeks (2001) permanent income hypothesis where there is no directly observable measure of permanent income. In practice, the existence of true measurements for theoretical counterparts is the exception to the rule.

Structural time series models that can handle unobserved components such as the business cycle, seasonality and trend has been developed by Harvey (1981) in time series analysis. More recently economists have analyzed the extent to which monetary policy is endogenously determined. However, given that there is no single measurable quantity which represents monetary policy, studies such as Avery (1979) have used a model specification where it assumed that monetary policy is represented by a single latent variable, and that policy is manifest in the behaviour of a set of indicators. In Lahiri (1976) the author examines the impact of the Fisher effect of inflationary expectations on the nominal interest rate, where the unobserved price expectations variables is modeled using a structural equations approach.

Casual models which incorporate latent variables have been utilized in several empirical studies; among others include Hauser (1972), Wolfe and Behrman (1984), Shehzad, (2004 and 2006), Giuffrida *et al*, (2005), Andreas and Fredriech (2008) and Christian and Manasa (2009).

Latent variables can have different definitions at different occasions and situations. To mention some there can be informal definitions, local independence definition, expected value definition, nondeterministic function of observed variables definition and sample realization definition[3]. In statistical models we can come across with different latent variables including regression disturbances as latent variables, latent variables in Limited-Dependent Variables models, latent variables in factor analysis, latent curve models, item response theory, latent class analysis and finally in structural equation modeling with latent variables which is the main point of application for this thesis[4].

The particular modeling strategy that is discussed in this paper has been extensively used in psychometrics and more recently in econometrics. It is founded up on the specification of a system of equations which specify the relationship between a set of unobservable latent

[3] For details about the definitions please see Bollen (2002)

[4] See Bollen for further discussions about the various latent variables properties under various circumstances in statistical models

variables, a set of observable endogenous indicators, and a set of observable exogenous variables. This approach builds up on the early work of Joreskog and Goldberger (1975) and Zellner (1970), and has been formalized in the LISREL[5] model of a set of linear structural equations. The modeling is presented in section 3.1 and 3.2.

2.5 Overview of structural equation modeling with Latent Variable

When analyzing the determinant of health, there are two main reasons why Structural Equation Models (SEM) with latent variables have recently been preferred to single equation models with a single health indicator as dependent variable. SEM enables a researcher to test a set of regression equations simultaneously. Thus, the main advantage of SEM is to construct a model that combines the determinants of different health states. Using this approach, one can consider simultaneously both the direct and indirect effects of variables such as income, age or education. For example, in the analysis of permanent and transitory health status it is important to separate the direct effect of some key variables from their indirect effect. However, it is not possible to disentangle these factors in a single equation model in which the reduced form parameters include both the direct and indirect effects. In addition, the central concept of health is inherently unobservable. Treating this factor as latent variable allows the inclusion of a variety of indicators that are partial measure of the underlying latent unobservable variables (Giuffrida *et al*, 2005).

Manning *et al*. (1982) showed that using multidimensional health status measures rather than a single item measure (e.g. self-assessed health status) increases the precision of the estimation significantly. The main techniques used to estimate structural equation models with latent variables are the linear structural relationship (LISREL) model developed by

[5] LISREL is an acronym for ***LI****near* ***S****tructural* ***REL****ation*

Jöreskog and Sörbon (1987). In the health economics literature there are numerous applications of LISREL methodology.

The majority of studies that used LISREL methodology adopted the Multiple Causes-Multiple Indicators (MIMIC) model, which is a special case of LISREL with only one latent variable – health status and the other extensions. Robinson and Ferrara (1977) were the first to use this approach. They presented a ranking of the 50 US states for 1969 based on a "health index" estimated by a MIMIC model. Wolfe and Van der Gaag's (1981) study is one of the first published papers that used individual level data. They used data on children from the 1975 Rochester Community Child Health Survey to construct a MIMIC model of health status and health care utilization. Several other studies employed this methodology. For example, Wagstaff (1986; 1993) estimated empirical specifications of the Grossman's (1972) health production model using data from Denmark. Giuffrida *et al*, (2005) have also estimated a Structural Equation Modeling with three latent variables (health status, household wealth and health care utilization) using Brazilian household data

Häkkinen (1991) estimated a MIMIC model of health and health care utilization using survey data from Finland. Among studies that allowed for more latent variables, we have van de Ven and Van der Gaag (1982), who estimated a SEM where both health status and income were latent unobservable variables using a health care survey carried out in The Netherlands. They found that the two variables had a mutual, positive impact and that age and education had important impacts on health status. Ersland *et al.* (1995) estimated the impact of environmental pollution on health status and health care utilization treating both variables as latent. The model was estimated using data taken from the German socio-economic panel and the quality of the environment turned out to be an important determinant of health status.

However, there are few health applications of Structural Equation Modeling in developing countries and does not exist in Africa. Behrman and Wolfe (1984) analyzed health status and health care utilization of women and children in Nicaragua using data collected in 1977-78. The authors expressed doubts on the positive effect of maternal schooling on children health and nutrition. They also concluded that several important characteristics that are usually associated with development had no or had negative associations with women's health: women's schooling, household resources and women's labor force participation. Shehzad (2006) used MIMIC models to identify the determinants of child health status and found significant impact of parental education, household size and public environment variables to be significant variables using the Pakistan's demographic survey. Giuffrida *et al.*, (2005) estimated the structural equation models to identify the factors determining the latent variables of health status, household wealth and health care utilization using the Brazilian household survey.

2.6 The Social Impacts of Economic Policies: in light of children

Recognition of the social impacts of different types of economic policies on different socio economic groups, including children, is long overdue. The impacts on children of macro-economic policies are usually mediated through the family or other local institutions such as school and health services. Certain economic policies have profound impacts on children. Because the impacts of these policies are mediated through a range of social institutions there is some unpredictability involved (Harper and Marcus, 1999/00).

There is little study that directly investigates effect of economic policies on children. This may reflect in part the difficulties of doing so, but also the extent to which the costs of such policies to children have been ignored. This gap in analysis overlooked the fact that childhood is a once-and-for-all window of opportunity for biological and social development.

Harper and Marcus (1999/00) argued that policies which lead to reduction in pro-poor expenditure, increased economic stress, through high unemployment or falling incomes, even if intended as a relatively short term effect, have long term implications because of the lasting impact on children.

Mehrotra (1998) noted that expenditures on public services are insufficient to achieve universal access in the next 5 to 10 years particularly with rapid population growth. In the 1980's most adjustment programmes that have originated from the Breton woods institutions cut expenditure on basic services in order to achieve macroeconomic balance. Although current policy promotes greater investment in public services, their neglect in the 1980's and early 1990's has already jeopardized the development of at least a 10 year cohort of children.

It is also stressed that for children and for society as a whole economic policies which advocate "*short term pain for long term gain*" are pursuing a strategy which can impact very negatively on children and hence to the future generation. The period of "*short term pain*" could encompass the entire developmental years of a child who will never recoup those losses.

Chapter III

3. Methodology and Model Specification

3.1 Structural Equation Modeling: what is it all about?

The basic idea of statistical methods falling under the general title 'Structural Equation Modeling' is to analyze the covariance structure of variables observed in the sample. More specifically, to analyze the difference between the population (sample) covariances and the covariances predicted by the model. Formally the fundamental hypothesis is stated as (Bollen, 1989 and Kiiskinen, 2003)[6]:

$$\sum = \sum(\theta) \qquad (3.1)$$

Where $\sum$ denotes the matrix of known (observed) covariance in the population (sample) and θ is a vector of model parameters. The aim of the modeling is to minimize the difference between the population covariance matrix and the model covariance matrix so that the above equation holds with some acceptable statistical precision.

It is possible to derive a system of equations from the covariance structure of equation (3.1) by defining the type of variables and how they are assumed to be connected with each other from a theoretical point of view. In the models containing both observed and latent (unobserved) variables, equations can be divided in to two: *structural equations* and *measurement equations*. In the case where latent variables are all endogenous the structural equations part of the model can be written as;

---

[6] The symbols and formulas are presented in accordance with the notation of the statistical software package LISREL, PRELIS and SIMPLES.

$$\eta_1 = \mathbf{B}\eta_2 + \mathbf{\Gamma x} + \zeta \quad (3.2)$$

Where η_1 and η_2 is a vector of latent variables, $\mathbf{x}$ is a vector of exogenous variables, $\mathbf{B}$ is coefficient relating the latent variables, $\mathbf{\Gamma}$ is the coefficient relating the exogenous variables with the latent variables and ζ is a vector of error terms. In addition the measurement model for the latent variable η is defined as:

$$\mathbf{y} = \mathbf{\Lambda}_y\eta + \varepsilon \quad (3.3)$$

Where y denotes observed indicators for the latent variable, η represents the latent variable itself, $\mathbf{\Lambda}_y$ is the factor loadings of the latent variable to their indicators and the error term ε represents the measurement error associated with each observed indicator.

Equations (3.2) and (3.3) together with the models assumptions can be used to express the elements of $\sum(\theta)$ in terms of models parameters, which is called the implied covariance matrix. After substituting the estimated values of parameters the $\sum(\hat{\theta})$ is compared to the sample covariance matrix **S** (representing $\sum$). In order to analyze the difference $(S - \hat{\Sigma})$ an appropriate function to be minimized must be chosen (e.g. see Bollen, 1989 and Kiiskinen, 2003). The selection of the function depends on the preferences of the researcher as to the representativeness of the function to the model and the extent to which the function tolerates to existing problems in the data.[7]

Structural Equation Model is hence used as the main analytical technique in this paper. It has been frequently used in these types of studies and health production models and demand for health studies as was reviewed in chapter two. There are several advantages in adopting the

[7] See section 3.3 for a discussion on this particular issue

SEM technique that are particularly useful considering the theoretical frameworks and the nature of data described in this chapter's section 3.9:

1. It can incorporate latent variables measured by multiple indicators such as those and hence is able to deal with underlying "theoretical variables".

2. It may conveniently be used to estimate a system of equations in structural form.

3. The use of latent variables and structural form estimation strengthens the link between the theory and the empirical analysis.

4. It utilizes all available information from a large number of variables in a simultaneous estimation framework.

Hence, SEM is used as the main analysis technique throughout the empirical analysis presented in Chapter 4 of the thesis using its special case MIMIC model discussed in the next section 3.2.

3.2 The MIMIC Model

The MIMIC model explains the relationship between observable variables and the unobservable variables by minimizing the distance between the sample covariance matrix and the covariance matrix predicted by the model. The observable variables are divided in to *causes* of the latent variables and the *indicators* (Joreskog and Goldberger, 1975; Andreas and Fredrick, 2008). Formally, the MIMIC model consists of two parts; the *structural model* and the *measurement model*. The structural equation model is given by;

$$\eta_i = \gamma' x_i + \varsigma_i, \quad \ldots\ldots (3.4)$$

Where $x'_i = (x_{1i}, x_{2i}, ..., x_{qi})$ a (1 X q) vector of variables as is indicated by the subscript i. Each $\mathbf{x}_i$, is a potential cause of latent variable η_i. $\gamma' = (\gamma_1, \gamma_2, ..., \gamma_q)$ is a (1x q) vector of coefficients in the structural model describing the "causal" relationship between the latent variable and its causes. Since the structural equation model only partially explains the latent variable η_i, the error term ς_i represents the unexplained component. The MIMIC model assumes that the variables are measured as deviations from their means and that the error term does not correlate to the causes, i.e. E (η_i) = E (ς_i) = 0 and

E ($x_i \varsigma'_i$) = E ($\varsigma_i x'_i$) = 0. The variance of ς_i is abbreviated by ψ and $\mathbf{\Phi}$ is the (q X q) covariance matrix of the causes x_i.

The measurement model represents the link between the latent variable and its indicators. i.e. the latent unobservable variable is expressed in terms of observable variables. It is specified by:

$$y_i = \lambda \eta_i + \varepsilon_i \quad \text{...(3.5)}$$

Where $y'_i = (y_{1i}, y_{2i}, ..., y_{pi})$ is a (1 X p) vector of indicator variables y_{ji} , j=1,…,p. ε_i=(ε_{1i}, ε_{2i},…,ε_{pt}) is a (p x 1) vector of disturbances where every ε_{ji} , j=1,…,p is a white noise error term. Their (p X p) covariance matrix is given by Θ_ε. The single λ_j, j=1,…, p in the (p X 1) vector of regression coefficients λ, represents the magnitude of the expected change of the respective indicator for a unit change in the latent variable.

It is assumed that the error terms in the measurement model do not correlate either to the causes x_i or to the latent variables η_i, hence,

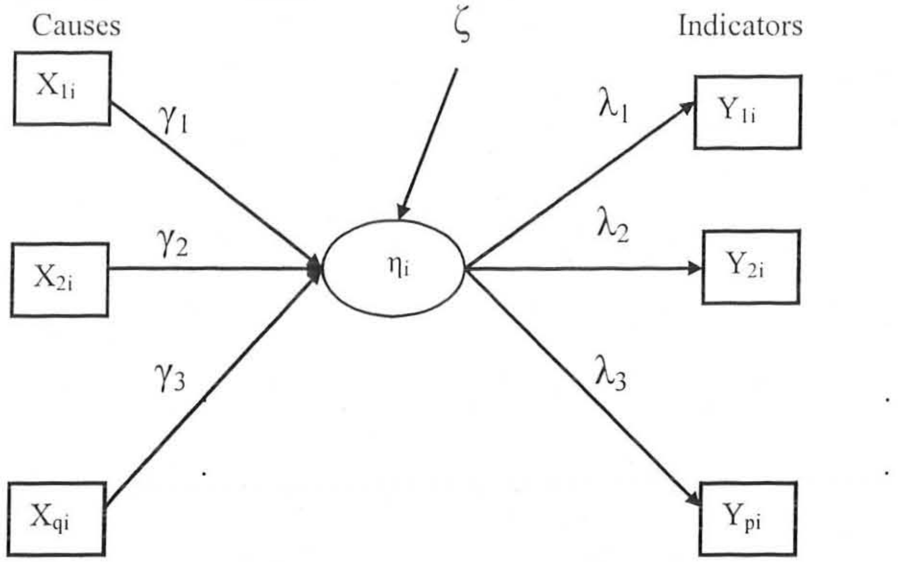
E ($x_i \varepsilon'_i$) = E ($\varepsilon_i x'_i$) = 0 and E ($\eta_i \varepsilon'_i$) = E ($\varepsilon_i \eta'_i$) = 0.

A final assumption is that the ε_i's do not correlate to ς_i, i.e.

E ($\varepsilon_i \varsigma'_i$) = E ($\varsigma_i \varepsilon'_i$) = 0.

Figure 3.1 shows the general structure of the MIMIC model.

Figure 3.1: A general structure of a MIMIC model



From equation (3.4) and (3.5) and making use of the definitions, we can derive the MIMIC model's covariance matrix Σ (see Appendix A). This matrix describes the relationship between the observed variables in terms of their covariance. Decomposing the matrix derives the structure between the observed variables and the latent variables, here the child health status. The model's covariance matrix is given by:

$$\Sigma = \begin{pmatrix} \lambda(\gamma'\Phi\gamma + \Psi) + \Theta_\varepsilon & \lambda\gamma'\Phi \\ \Phi\gamma\lambda' & \Phi \end{pmatrix}$$

Where Σ is a function of the parameters λ, γ and the variances contained in Φ, Θ_ε, and ψ. Since the latent variable is not observable, its size is unknown, and the parameters of the model must be estimated using the links between the observed variables' variances and covariance. Thus, the goal of the estimation procedure is to find values for the parameters that produce an estimate for Σ that is as close as possible to the sample covariance matrix for the observed causes and indicators, i.e. the x_i's and y_i's.

Such models provide possibilities to model more complex systems with several jointly dependent variables. Also, they provide the possibility to model health as an unobservable variable that is determined by a set of its causes and indicators, which by themselves would each constitute only a partial measure of health. Moreover, unobservables other than health may be included in the model (Van de Ven and Van der Gaag, 1982) if the data set provides an adequate number of useful variables. One of the advantages of MIMIC-modeling is the possibility to estimate the whole system of equations simultaneously and hence to test the specification of the entire model using a single test statistic (Hakkinen, 1991).

3.3 Estimator

As mentioned previously, a strong preference has to be given to the modeling techniques and the estimation methods that account for the endogenous nature of individual choice variables, and, moreover, rule out the influence of any unobservable confounding factors.

The structural equation models such as MIMIC require estimation techniques that are based on more restrictive assumptions than simultaneous equation methods such as 2SLS. Estimation does not allow for possible unobservable influences outside the system to be considered. The limitation of the FIML (Full Information Maximum Likelihood) estimators for such models is that the assumption of multivariate normality may not hold if discrete health indicators are used and even if the consistency of the parameter estimates is maintained the corresponding standard errors are not valid. In order to ensure the identification of the model several parameters have to be usually constrained a priori, which may also prove not to be supported by the data (e.g. Van de Ven and Van der Gaag, 1982).

Due to the non-ideal characteristics of data the use of generally Weighted Least Squares (WLS) is considered as the primary estimation method. The general form of the WLS fitting function is:

$$F(\theta) = (s - \sigma)' \mathbf{W}^{-1}(s - \sigma), \quad \ldots\ldots(3.7)$$

Where F refers to function, 's' is a vector containing the elements on and below the diagonal of the sample covariance matrix **S** and σ is the vector of the corresponding element predicted by the model. Since, in reality, the correlation structure P (θ) rather than covariances will be fitted to the correlation matrix the more consistent expression of (3.7) would be (Joreskog *et al.*, 1999):

$$F(\theta) = (r - \rho)' \mathbf{W}^{-1}(r - \rho), \quad \ldots\ldots(3.8)$$

Where the observed sample correlation r and model correlation ρ are used instead of s and σ, respectively. The weight matrix **W** is the asymptotic covariance matrix, which, in the case of (3.8), is based on correlations instead of covariances in case of (3.7). The parameter vector θ in (3.8) has (p + q) elements less than in (3.7) since the diagonal elements of P (θ) are fixed to unity (Joreskog *et al.*, 1999).

Even though the parameter estimates of other fitting functions (especially ML) are quite robust towards the non normality and appropriate corrective procedures for standard errors and chi-square statistics exist, fitting the correlation structure to the correlation matrix requires the use of WLS (Joreskog *et al.*, 1999).

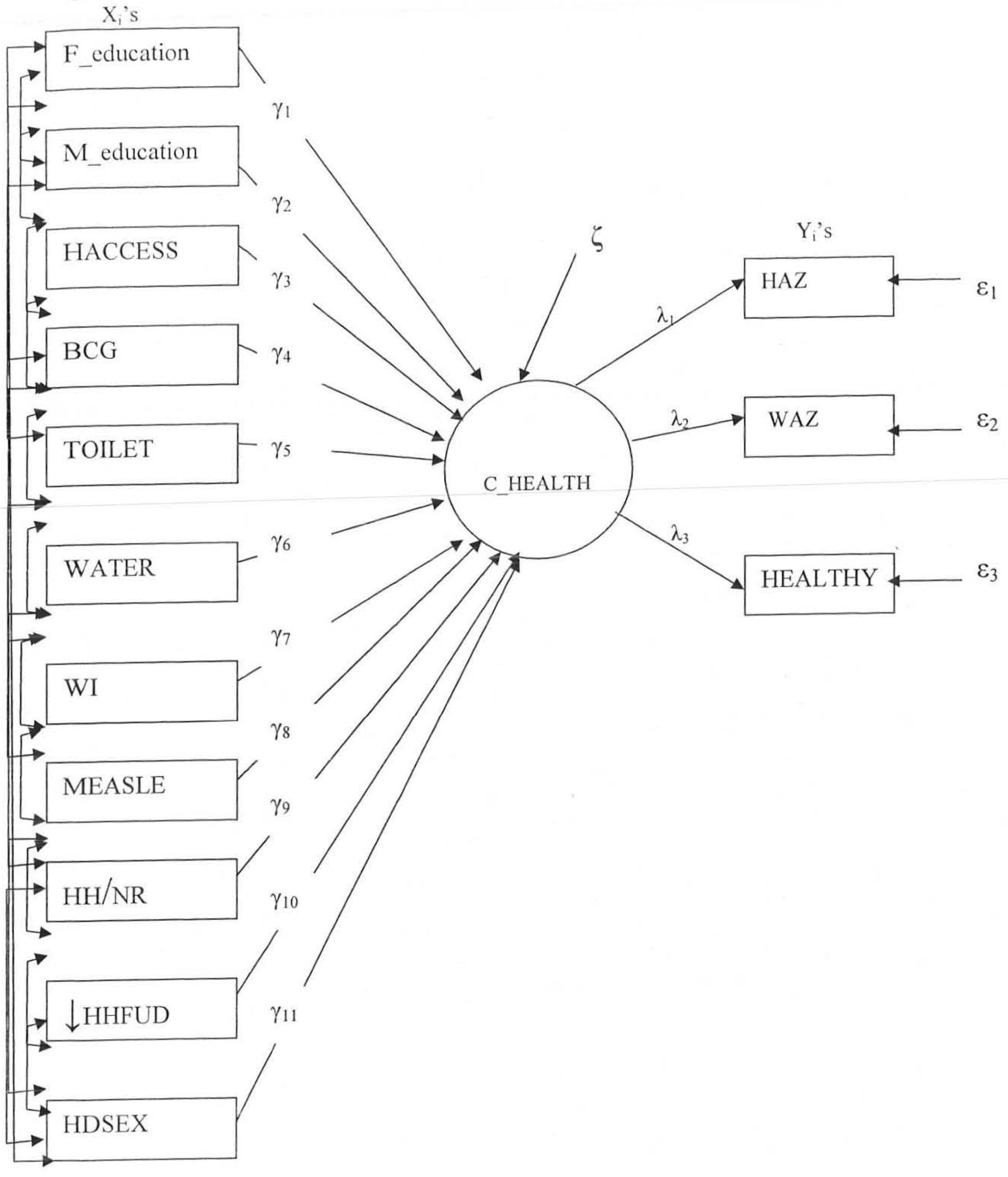
3.4 Empirical Specification: to determinants of child health

The empirical specification of the model to child health followed two steps for illustrative purposes; it starts with single latent variable and proceeds to two latent variables adding interaction effect for the last case. Since the objective of the study is to look at the factors behind the latent variable child health status the study has given emphasis to the last case of latent variables (See Shehzad ,2006).

3.4.1 Single latent variable

MIMIC (Multiple Indicators Multiple Causes) models estimates the effects of multiple explanatory variables on the latent variable, child's health state. The model can best be understood by means of a path diagram for one latent variable case. In Figure 3.1, the unobservable 'η' variable, child health status is enclosed in a circle and the observed variables both the indicator variables (y_i's) to the right of the latent variable and the explanatory variables (x_i's) to the left are represented in the rectangles. The error terms are not enclosed.

Figure 3.2: Path diagrams for single latent variable[8]



[8] See section 3.9's Table 3.1 for the definition of variables with in this Figure 3.1

The arrows leading from x to the latent variable indicate the hypothesized direct effect on child health status. The arrows leading from the latent variable, child health status, to the indicators on the right hand side represent hypothesized impact of unobservable variable on the indicators. The strength of the effects of the variables is indicated by coefficients in the measurement model. Using the LISREL terminology, the model is as follows.

A. Measurement model

$$y = \lambda_y \eta + \varepsilon \qquad (3.9)$$

Where y is the indicator of the latent variable, η is the latent variable, λ_y is the factor loadings from the latent variable to the indicators and ε is the error term in the measurement model.

B. Structural Equation Model

$$\eta = \gamma' \mathbf{x} + \zeta \qquad (3.10)$$

, where η refers to the latent variables as above, **x** refers to the exogenous variables in the model, γ is the coefficient of the variables showing the strength of the impact of the exogenous variables on the latent variable and ζ is the error term in the structural model.

The following are the assumptions for the MIMIC model with single latent variable;

$E(\eta) = E(\varepsilon) = 0$ $\qquad$ $E(\zeta) = 0$

$E(\varepsilon, \eta) = 0$ $\qquad$ $E(\zeta, \varepsilon) = 0$

$E(\zeta, \eta) = 0$

3.4.2 Two latent variables

The path diagram representation for the two latent variables case is presented in Figure 3.2. The specification of the model for the latent variables is as follows using the LISREL terminology;

A. Measurement model

$$Y=\lambda_y\eta + \varepsilon \qquad (3.11)$$

Where y is the indicator of the latent variables, η is the latent variable itself, λ_y is the factor loadings of the latent variables to the indicators and ε is the error term in the measurement model.

B. Structural Equation Model

$$\eta = B\,\eta + \gamma\mathbf{x} + \zeta \qquad (3.12)$$

The following are the matrices/vectors specified in the model

1. λ_y (p x m) is the matrix of coefficients, or loading, relating indicators of endogenous variables to latent exogenous variables (η).
2. γ (q x n) is the matrix of coefficients, or loading, relating indicators of exogenous variables to latent exogenous variables (ξ)
3. B (m x m) is a matrix of coefficients of the effects of latent endogenous variables on latent endogenous variables.
4. Γ (m x n) is the matrix of coefficients of the effects of latent exogenous variables on latent endogenous variables.
5. Φ (n x n) is a variance-covariance matrix of the latent exogenous variable (ξ)
6. ψ (m x m) is a variance-covariance matrix of the residuals (ξ)

7. Φ_{ε} (p x p) is a variance-covariance matrix of errors of measurements of y's

8. Φ_{δ} (q x 1) is a variance –covariance matrix of errors of measurement of x's

9. η (m x 1) is a vector of latent variables.

10. ζ (m x 1) is a vector of error terms in the structural model.

11. **y** (p x 1) is a vector of indicators for the latent variable.

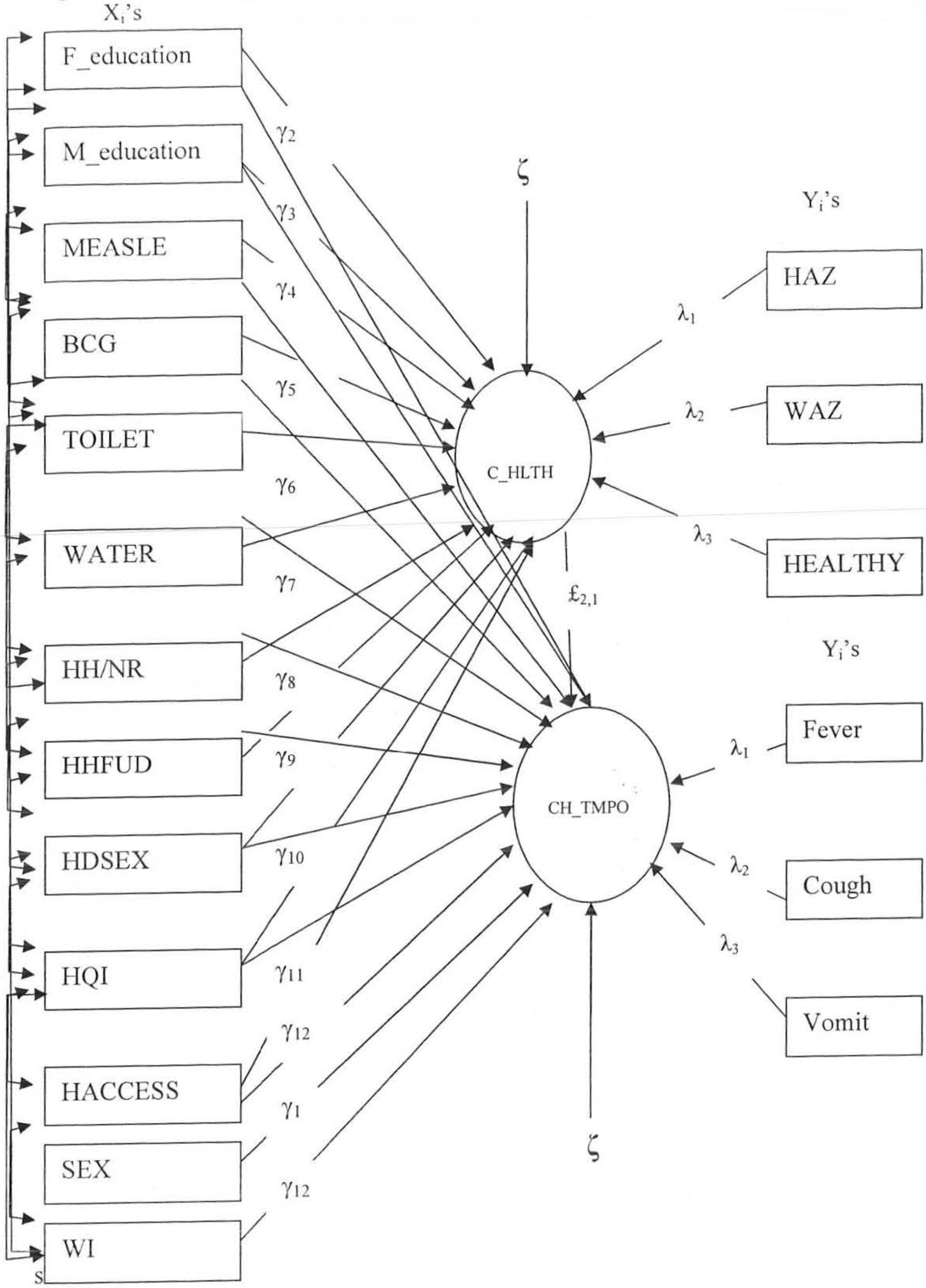
12. ε (p x 1) is a vector of error terms in the measurement model.

13. **x** (n x 1) is a vector of exogenous variables.

Here p is the number of y variables; m is the number of η variables, n is the number of ξ variables and q is the number of x variables. The path diagram for two latent variables case is presented below[9].

[9] Please see section 3.9 for a discussion about the variables and reason for selection.

Figure 3.3: Path diagram for two latent



3.5 Testing the Structural Equation Models

The fundamental question in evaluating a SEM model has traditionally been whether the specified model results in an exact reproduction of the population covariance matrix (Kiiskinen, 2003). For a correctly specified model and data with appropriate distributional properties an asymptotically χ^2 distributed test statistic could be used to test the null hypothesis, represented by the Equation (3.1).

The test statistic is derived from the minimum value of the fitting function:

$$c = (N-1)F(S, \Sigma(\hat{\theta})) \qquad (3.13)$$

Where F(S, $\Sigma(\hat{\theta})$) is the estimated (minimum) value of the fit function and N is the sample size. The function is approximately χ^2 distributed with d = k − t degrees of freedom in a large sample. k = (p + q) (p + q + 1)/2 is the total number of distinct elements in S, t is the number of free parameters estimated, and (p + q) is the number of observed variables. In principle, the model is rejected if c exceeds the (1 −α) percentile of χ^2-distribution (Joreskog and Sorbom, 1993a). In practice, this test is used more as guidance in modifying the model ("model generating approach") rather than as a strict rule for model rejection.

One of the problems with the χ^2 statistic is that it is affected by several factors, the most obvious of which is the sample size. If N is small the test may not have adequate power to detect the meaningful differences between $\hat{\Sigma}$ and S, whereas large N may lead to over rejection of the null hypothesis due to detection of minor discrepancies (Fan *et al.*, 1999). Other sources of variation in fit statistics, χ^2 and other measures based on it, include model complexity (degrees of freedom), sampling variation, selection of alternative models (if any), the estimation method, and whether a correlation or covariance structure is analysed (Tanaka, 1993; Bollen, 1989; Kiiskinen, 2003).

There are a vast number of measures available to assess the overall fit of the model. Basically, all of these statistics are derived from the minimum value of the fit function. Since chi-square is defined in a way that depends on the sample size, it tends to yield high values in large samples like in this study. In order to provide alternative ways of assessing model fit, a number of fit indices are available. Fan *et al.* (1999) have categorized these indices into four types:

1. Covariance matrix reproduction indices attempt to evaluate what proportion of the variation the model generated covariance matrix can explain the sample covariance matrix. For instance, the goodness-of-fit (GFI)[10] index fall into this category:

$$GFI = \frac{F[S,\Sigma(\hat{\theta})]}{F[S,\Sigma(0)]}, \qquad (3.14)$$

, which is based on a ratio of the estimated (minimum) value of the fit function to the value of the fit function where all parameters are zeros.

And within this category we have the Adjusted Goodness-of-Fit Index (AGFI):

$$AGFI = 1 - \frac{k(k+1)}{2d(1-GFI)} \qquad (3.15)$$

[10] See section 3.6 for the assessment criterion.

, which makes an additional adjustment for the degrees of freedom and naturally also belongs to this category.

2. The second category, comparative fit indices, are designed to assess model fit in comparison with the more restricted null model, usually the independence model (with no relations among indicators). Measures of this type have been discredited by the fact that the independence model has usually an extremely poor fit, and hence may not be an alternative specification in the first place (Fan *et al.*, 1999).

3. Finally, a third category of indices are based on the non-centrality statistics, and hence are less sensitive to moderate misspecification of the model. The comparative fit index (CFI) belongs to this category using the non-central χ^2 statistic. First, define $\hat{F}$ as minimum fit function value for the estimated model and as the corresponding value for the independence model, then the non-central statistic is

$$\tau = \mathrm{Max}\ \{(N-1)\ \hat{F} - d,\ 0\}$$

and

$$\tau_i = \mathrm{Max}\ \{(N-1)\ \hat{F}_i - d_i,\ (N-1)\ \hat{F} - d,\ 0\} \quad \text{.....(3.16)}$$

, for the estimated and independence models, respectively. d and d_i denote the corresponding degrees of freedom. The CFI is:

$$\mathrm{CFI} = 1 - \tau / \tau_i \quad \text{.....(3.17)}$$

Yet another approach to model evaluation is provided by the Root Mean Squared Error of Approximation (RMSEA) which has become one of the standard fit measures to be reported with SEM models. It is based on population discrepancy function ($\hat{F}_o$):

$$\hat{F}_o = \text{Max}\ \{ \hat{F} - (d/(N-1)), 0\}, \ldots\ldots\ldots\ldots\ldots\ldots(3.18)$$

, where $\hat{F}_o$ approaches $\hat{F}$ when N approaches infinity, but decreases when parameter are added to the estimation with given N. In order to account for the variation in the degrees of freedom the RMSEA is defined as population discrepancy per degrees of freedom:

$$RMSEA = \sqrt{\hat{F}_o / d} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.19)$$

The RMSEA is considered to indicate very good fit if the value falls below 0.05 and a reasonable fit with values less than 0.10, whereas values greater than 0.10 would suggest some problems in fit (Fan *et al.*, 1999).

The RMSEA will be used as the main indicator for model fit since it has been found to be sensitive to model misspecification, but is still minimally dependent on sample size (Fan *et al.*, 1999). Also the CFI appears to be robust to variation in sample size and will be used as a secondary indicator of fit. Also, the multivariate counterpart of coefficient of determination (R^2) AGFI is included, but it should be interpreted more cautiously since it is known to be relatively sensitive to the sample size.

3.6 Model Assessment

The table below provides the model goodness-of-fit statistics that are applicable to MIMIC model.

Table 3.1: Model Fit Assessment and Test Statistics: Test Statistic Acceptance Criteria (Laurence H. Lester (2008))

Test Statistics	Purpose	Acceptance Criteria
Root Mean-Square Error of Approximation (RMSEA)	Absolute fit (0 is perfect fit <0.01 is outstanding)	< 0.05 close <0.08 good < 0.10 reasonable
Standardized Root Mean-Square Residual (SRMR)	Absolute fit	<0.10 favorable <0.05 good
Goodness-of-fit (GFI) Adjusted Goodness of Fit Index (AGFI)	Absolute fit (range 0 no fit, 1 perfect fit)	>0.90 good fit
Comparative Fit Index (CFI)	Incremental fit (range 0 to 1)	>0.90 good fit
Parsimony-based GFI (PGFI) and Parsimony-based Normed Fit Index (PNFI)	Incremental, parsimony adjusted, fit (range 0 to 1)	No defined level
Akaike Information Criterion (AIC) and Consistent Akaike Information Criterion (CAIC)	Comparative model fit (no upper limit, 0 perfect fit)	No defined level
Expected Value of the Cross-validation index (ECVI)	Comparative model fit (no upper limit, 0 perfect fit)	No defined level
Chi-squared	model fit (see note 1)	Significant => model fits the data

Notes: (1) The Chi-squared statistic is not a reliable goodness-of-fit indicator in large samples, but it is useful to assess the relative fit of various models—the Chi-squared statistic is the Satorra-Bentler Scaled Chi-squared, which takes non-normality of input data into account.

Models are based on analysis of the correlation matrix—goodness-of-fit is an assessment of how well the derived model replicates the observed correlation matrix—and, as there is no single goodness-of-fit measure, it is practice in applied work to report several appropriate statistics.

3.7 Evaluating the MIMIC Model Method

The information that comes out of models in this paper requires the application of advanced analytical techniques based on latent variables. Thus, an obvious question can be asked to evaluate the research outputs in this paper:

Consider the following relating to the advantages of the linked MIMIC models:

(1) Why not use a single measure to represent health in a common regression model? Measuring health by a single imprecise indicator variable (e.g. Anthropometric measures), fails to incorporate measurement error, and is a coarse measure that fails to incorporate the multidimensional aspects of health. That is, no single indicator sufficiently captures the multiple aspects of health and while it is not possible to capture all aspects of child health, multiple indicators cover multiple attributes.

(2) Is cross-sectional factor analysis (FA) appropriate? Cross-sectional analysis does not track dynamics of the child health. As with the benefits of panel data econometric models versus cross-sectional analysis, longitudinal data can examine change, differentiate between change in cohorts and change for individuals and reduce the possibility that a particular cross-section is a typical. Moreover, for survey data, across-time correlations distort empirical estimates if ignored. Thus the thesis work recommends the application of SEM to panel data for further research.

(3) Why not use the reduced form of the MIMIC model (i.e. remove the latent variable for child health and represent the underlying model in the more familiar econometric form)? It is well known that the MIMIC model can be re-specified as the reduced form in which only observed variables (i.e. formative and reflective indicators) appear. This allows estimation (as a system of simultaneous equations) but the path coefficient for the

formative (causal) indicators and reflective indicators cannot be extracted—only a composite coefficient is available, unless it is assumed that all variables are measured without error—and thus nothing can be said about the relative importance of reflective and formative indicators. As well as failing to incorporate measurement error, the reduced form system, (i) ignores across-time correlation between reflective indicator errors, and (ii) does not model correlation between formative indicators within and across time. Thus, the MIMIC model is not a standard (linear) simultaneous equation model, but a model which incorporates errors-in-variables, and associated correlations (see Jöreskog and Goldberger 1975).

In summary, although somewhat complex, MIMIC models provide an analytical method that appropriately deals with many important issues that are not dealt with in less sophisticated methods.

3.8 Interpreting the MIMIC Model: *Variance Explained—Model Reliability*

The reliability of the unobserved latent construct, child health, can be assessed by considering the implied proportion of variance in child health that is explained by the model—that is, the reliability measures how well the reflective indicators serve as instruments for child health (Jöreskog and Sörbom 2001).[11]

The measure of reliability shows that the proportion of variance explained varies, but lower values are not unexpected given that much of the data are dichotomous or ordinal, the exploratory nature of the models, and the opportunistic use of the data (which was not collected with the intention of applying sophisticated statistical techniques). On the other hand, higher reliability values suggest a remarkable degree of accuracy given these conditions.

[11] Child health is unobserved, but an estimate ***of its variance*** is a model output (which does not require an estimate of the latent variable itself).

In interpreting the MIMIC measurement model the path coefficients (factor loadings, λ,) are the effect of the latent variable on its reflective indicator (i.e. the extent to which ***Child health*** status is jointly reflected in ***WAZ, HAZ and HEALTHY***)[12]. And the path coefficients from the exogenous variable to the latents show the strength of the impact of the exogenous variable on the latent variable.

3.9 Data Source and Variables definition.

The source of data for this research is the Young Lives. The Young Lives is an international innovative longitudinal study of childhood poverty in Ethiopia, India, Peru and Vietnam. Between 2002 and 2015, some 2000 children in each country are being tracked and surveyed at 3-4 year intervals from when they are 1 until 14 years of age. In Ethiopia, the project has received financial support from the UK Department for International Development.

The study used quantitative data collected 2002 from 1,999 households with 6–18- month-old children from five regions of Ethiopia as a result of the Young Lives Ethiopia project. The data was collected from 20 sentinel sites with 100 children from each site. Eight sites are in urban areas and 12 are rural. They are mainly in food insecure areas in Addis Ababa and the Amhara, Oromia, Southern Nations, Nationalities and Peoples (SNNP) and Tigray regions. In the survey data, the poor were deliberately over-represented by purposive selection of sites as the objective of the project was to analyze panel data focusing on child poverty (Tekie A. *et al*, 2003).

[12] See LISREL 8.7: User's Guide; measurement model output, Scientific Software International, Inc available at www.ssiscience.com

Indicator Variables/Dependent Variables:

To start with the indicators used for the latent variables, the Young Lives Survey collected information on many indicators of child health status. Based on the conceptual framework used the study selected initially three dimensions of child health status. First, a measure of the care-giver perceived health status compared to peers of same age has been taken. This variable takes a value of one when the health status is same or better and zero otherwise. Second and third are the Z-Scores of the Height-for-Age and Weight-for-Age. The Z-scores are standardized using age sex specific median from the U.S. National Center for Health Statistics (NCHS) as recommended by the World Health Organization (WHO).

For the transitory health state measure that indicate temporary health disorder were selected; dummy for cough in the last 24 hours, dummy for vomiting everything in the last 24 hours and dummy for having high fever in the last 24 hours.

Explanatory Variables:

From the independent explanatory variables selected includes individual specific, household characteristics and community characteristics variables. The variables were selected based on the theoretical framework, their importance in the child health literature and their availability in the data.

From the individual specific variables the variable showing the gender of the child is included in the model. This variable takes the value of one for male and zero for female.

From parental and household characteristics many variables were selected including the level of education of the mother and father of the child, the ratio of household size to the number of rooms in the household, decrease in the availability of food in the household, whether the household is being headed by female or male etc. The ratio of household size to the number of rooms in the household (NNHR) is meant to capture the impact of higher family size on the health status of children in their first year of life. The educational background of both the father and the mother are taken to show the extent to which the parents level of education affect the health status of children. The dummy for father's education (F_EDUCAT, 1=Completed primary, 0=otherwise) indicates if the father had completed primary school and similarly a dummy for mother's education (M_EDUCAT, 1=Completed primary, 0=otherwise). This is because it is assumed that well educated parents take good care of their children, have good health state themselves, bear in mind the training and related information home for caring about their children, etc.

The dummy for a decrease in the availability of food in the household (HHFUD) is selected because we accept Wolfe and Behrman's (1981) idea that "*you are what you eat*". The other exogenous variable is the wealth index (WI) that shows the economic status of the household in which the YL child is living in. In addition to that, in developing countries assets that a household acquires can serve as indicators of the long run economic status of the household. The World Bank suggests the use of household consumer durables, household quality etc in measuring the relative economic position of a household (Ayalew, 2006). The indices used in this study are those constructed by the YL team. [13]

From the community variables include the dummy for the source of household drinking water, type of toilet and whether the child had been vaccinated for either BCG or measles. In order to look at the impact of economic policies on children the study selected those

[13] For the method used and variables included in the indices see Tekie A. *et al.*, (2003) page 23.

community variables that can represent the extent of government's infrastructural investment in the living areas of the children.

The first variable that indicates a dummy for the availability of pure drinking water supply infrastructure in the community is expected to have a profound impact on the health status of an average Ethiopian child. The second variable is a dummy for measles and a dummy for BCG vaccination each indicating the proximity of the household to a health center and also the health knowledge of the parents. If the household is closer to at least one health center it would most probably get the measles vaccination which is vital to the health status of children. Increasing the availability of these service centers is assumed in this paper to be the responsibility of the government. Thus its availability indicates the pro-children policy of the government.

Table 3.2: Definition of variables

Indicator Variables	Definition	Category labels
Indicator Variables		
WAZ	Weight-for-Age z-score	Amount
HAZ	Height-for-Age z-score	Amount
HEALTHY	Child health compare to others	1=same or better 0=otherwise
FEVER	High fever in last 24 hours	1=yes and 0 No
COUGH	Cough in last 24 hours	1=yes and 0 No
VOMIT	Vomit in last 24 hours	1=yes and 0 No
Explanatory Variables		
SEX	Gender of the child	1=Male and 0= Female
HHNR	Ratio of household size to the number of rooms in the household	-
F_EDUCAT	Fathers level of education	1=completed primary and 0=otherwise
M_EDUCAT	Mother's level of education	1=completed primary and 0=otherwise
MEASLE	Had Measles Vaccination	1=yes and 0 No
BCG	Had BCG Vaccination	1=yes and 0 No
HHFUD	Decrease in food availability	1=yes and 0 No
HDSEX	Gender of household head	1=Male and 0=Female
WI	Wealth Index	amount
WATER	Main source of drinking water	1= if safe/protected, 0=otherwise
TOILET	Type of toilet facility	1= pit latrine or flush, 0=otherwise
HACCESS	Distance from Nearest Health care center	1= if less than 10 km 0=otherwise

Children in this study are defined as those individuals of one year old as they are those vulnerable to health status deviations from both permanent and transitory health and because the study aimed to look at early childhood health. Moreover, the first year of a child's life is decisive for the child's later health status. Therefore, the data on this early year allows us to analyze factors affecting the foundation of health status.

The summarized major characteristics of all the variables used in this study are presented in Table 3.3 below with their respective means, standard deviation, skewness and kurtosis.

Table 3.3: Summarized descriptive statistics of all variables

Variable	Mean	St. Dev	Skewness	Kurtosis
Dummy for Sex	0.521	0.500	-0.084	-1.995
HH to No. of Rooms	4.503	2.247	0.836	0.578
Dummy for Father's Education	0.258	0.438	1.108	-0.774
Dummy for Mother's Education	0.325	0.469	0.748	-1.443
Dummy for Source of Drinking Water	0.126	0.332	2.250	3.068
Dummy for Comparative Health	0.755	0.430	-1.187	0.000
Weight for Age	-1.681	1.389	0.174	0.381
Height for Age	-1.516	1.719	0.438	1.419
Dummy for Measles Vaccination	0.675	0.469	-0.748	-1.443
Dummy for Decrease in Food Availability	0.325	0.469	0.000	-1.443
Wealth Index	0.180	8.182	5.505	1.479
Dummy for Healthcare Access	0.532	0.131	-7.367	0.000
Dummy for FEVER	0.267	0.444	1.071	-0.797
Dummy for COUGH	0.095	4.024	-23.828	573.383
Dummy for VOMIT	0.025	3.280	-29.247	864.081
Dummy for BCG vaccination	0.264	0.4413	-0.863	1.066
Dummy for Toilet facility	0.379	0.485	-1.753	0.4983
Dummy for Household Sex	0.141	0.348	2.233	2.057

Chapter IV

4. Estimation Results and Interpretation

4.1 Determinants of Child Health Status

In developing a complex structural model corresponding to the theoretical framework represented in the earlier chapter, valid measurement models are prerequisite to the analysis of structural relationships between variables of interest. Three indicators are chosen to best describe the latent concept of health status. The model's result can best be presented and understood by means of a path diagram that is presented in Appendix C. The estimation result of the single latent variable MIMIC models output is summarized in the following table;

Table 4.1: MIMIC Model for one latent variable child health status.

VARIABLES	C_HEALTH	$\hat{\rho}$#
Measurement model	*No. of Observation = 1845*	
Weight-for-Age Z-score	*0.033**** *(4.692)*	*0.13*
Height-for-Age Z-score	*1.000*	*0.40*
Dummy for Comparative Health (1 if same or better)	*0.176**** *(5.696)*	*0.52*
Structural models		
Ratio of Household size to number of rooms in the household	-0.083* (-1.551)	
Dummy for Father's Education (1 if completed primary)	0.447*** (6.641)	
Dummy for Mother's Education (1 if completed primary)	0.326*** (3.562)	
Dummy for Measles Vaccination (1 if yes)	0.028 (0.167)	
Dummy for Decrease in food Availability (1 if yes)	-0.204* (-1.822)	
Dummy for Sex of the Household Head (1 if male)	0.015 (1.283)	
Dummy for Source of Drinking water (1 if safe/protected)	0.137** (2.803)	
Dummy for Toilet Facility (1 if flush or pit latrine)	2.428*** (15.435)	
Wealth Index	0.278*** (5.332)	
Test Statistics[14]		
χ^2	300.265 DF=20 p=0.0000	
AGFI(Adjusted Goodness of fit Index	0.997	
RMSEA (Root Mean Squared Error Approx.)	0.084	
SRMR(Standardized Root Mean-Square Residual)	0.046	
ECVI(Expected cross-validation Index)	0.201	

Note: the values in parenthesis are t-values and where the t-values are not reported the indicators are set equal to one *** represents significance at 99%, ** represents significance at 95% and * represents significance at 90%. # refers to the squared multiple correlations.

[14] Interpretation and assessment criterion is presented in section 3.6 and 3.8

The above Table 4.1 and the path diagram in Appendix C depict the MIMIC models result for child health containing both the measurement model and the structural model. The selected indicators for the latent variable child health are the *weight-for-age z-score*, the *height-for-age z-scores* and the *care-giver perceived relative health status* of the child as compared with peers of same age. The factor loading of the variable *Height-for-Age z-score* has been set to be equal 1 to assign a unit of measurement for the latent construct and avoid the problem of indefiniteness. Hence this variable and the latent construct are set to be equal.

The Table 4.1 above also shows the overall fit of the model. The test results show that the overall fit of the model is very good as indicated by Adjusted Goodness-of-Fit Index (AGFI), Root Mean Squared Residual (RMR), and the significance of χ^2, that the overall fit of the models is very good. A very high AGFI of about 99 percent, a lower Expected Cross-Validation Index (ECVI) of about 0.221 and significant χ^2 and low SRMR of about 0.046 suggest that the model fits the data very well.

The squared multiple correlation ($\hat{\rho}$) which is also interpreted as a measure of the reliability of each indicator, are rather modest. The highest $\hat{\rho}$ is 0.52 for the *dummy of comparative health* indicates that 52 percent of the variation in care-giver perceived child health status is explained by latent health status variable, which is the permanent health state. The lowest $\hat{\rho}$ is only 0.13 for *Weight-for-Age z-score*. Other studies have also reported problems in trying fit a single latent variable model for health related behaviors (Boniface and Tefft, 1997; Kiiskinen, 2003), and it may well be the case that more than one dimension must be included in such models. However, the model is already relatively large, and therefore it is not a desirable strategy to increase the number of observed variables in this part of the model in order to allow for the multidimensional nature of factors affecting child health.

However, the t-statistics shows that all of the indicators selected are very significant at 99 percent confidence level. The model reliability estimate Root Mean-Squared Error of Approximation (RMSEA) shows in the above Table 4.1 that the model is reasonable to the data with 0.084. The t- statistics also shows that except the *dummy for Measles Vaccination* (1 if yes) and *dummy for Head of the Household* (1 if male), all the rest of the independent exogenous variables are significant factors that determine the level of child health status. We organized our discussion of these estimates as follows.

Household/Individual characteristics

The study finds evidence that there is a wide spread negative impact of an increase in the *ration of household size to number of rooms* in the household on child health. This implies that the higher the number of people sharing one room in the household the lower the health status of the child. This may result from *congestion effects* resulting in health disorder and easy transmission of diseases among the household members. This result is consistent with many other studies among others include Alderman (1990); Sentayehu (2004); Shehzad (2006) and Behrman and Wolfe (1981) who showed that an increase in household size has a negative relation with child health and nutrition.

The variables indicating the age and sex of the child are not incorporated in the model because the indicators of the latent construct are gender and age standardized.

The dummy variable representing the *sex of the head of the household* is found to have a positive relationship with the health status of children, but it is insignificant. This implies, at least from the relationship that in those households headed by male the health status of children is higher. This is may be because of the division of labor that exists in the household for work outside the household and home care for children.

Parental education

The model also has shown that mother's education has a significant positive relation with the health status of the child. This implies that a child whose mother completed at least primary school has a higher health status than others. The main reason as justified in the literature is that, child-care is the primary responsibility of the mother, whereas, fathers are responsible for breadwinning and earning. Moreover, more educated mothers may have healthier children because they have better knowledge about health care and nutrition, have healthier behavior, and provide more sanitary and safer environments for their children (Behrman and Deolalikar, 1988, 1990; Strauss, 1990; Thomas *et al.*, 1990, 1991; Desai and Alva, 1998; Glewwe, 1999; Currie and Moretti, 2003).

In addition to the nurturing effect, nature could also play an important role. More educated mothers are more likely to have better health, which genetically leads to better health for their children (Behrman and Wolfe, 1987; Wolfe and Behrman, 1987). The identification of mother's education as a major determinant of child health status enables us to evaluate a development policy of parental education. The result of this paper in relation to mother's education is consistent with the findings of other studies such as Alemu *et al.* (2005b); Chen and Li (2009); Christiaensen and Alderman (2001); Barrera (1990); Thomas *et al.* (1990a); Senaur and Garcia (1991); Escobal (2005); Shehzad (2006); Silva (2005); Woldemariam and Timotios (2002).

However, it has been also proved in other studies same us ours (Shehzad, 2006 and Behrman and Wolfe, 1981) that an increment in father's education has a positive effect on children's health status. Therefore, it can be said that an emphasis on parental education for improvement in child health should be integrated in the overall health system development policy.

Household Wealth

The *wealth index* which is an indicator of household economic status appears to have a substantial impact on child health which is significant at 1 percent significance level. Specifically, an improvement in the *wealth index* by just a unit brings 0.28 units improvement in the health status of children. Therefore, we conclude that our estimates are consistent with there being an important influence of the household economic status on the health status of a child. In other words, it means that as the economic status of households gets better, the health status of a child improves. The same result has been obtained for the *household quality index* with a lesser significance level as compared to the *wealth index*.

Our result shows that at 10 percent level of significance a decrease in the household food availability has a negative consequence on the health status of children which is consistent with our prior expectation. Thus, this implicates to the fact that variability in the availability of food in the household makes the health status of children volatile throughout their lives. We have found thus a support to our earlier proposition that "*you are what you eat*". The result is consistent with that of Behrman and Wolfe (1980); Dercon and Hoddinot (2003) and Alemu *et al.* (2005). Therefore, the elimination of factors which make the food availability at home volatile has to be reduced if an increased level of child health is to be maintained. For instance, since most people in Ethiopia depend on agriculture for their livelihood an effort has to be made to eliminate all factors that make the level of agricultural output variable such as looking for ways of transforming the rain fed agriculture to an irrigation system.

Government Policy variables/Public Environments

The dummy variable on *Measles Vaccination*, though insignificant has the expected sign. This result implies that those who did not get vaccination with measles have low level of health status as compared with those vaccinated. Thus this implicates to the fact that if the government expands the availability of centers for vaccination service such as primary health centers in all areas, the level of child health status will be improved. In other words, all policies that support and advocate for the expansion of health extension service, the construction of health centers and increasing the number of trained health personnel thus have positive impact on child health. Any policy move against this, such as a shifting of budget to other priority areas will have a profound negative consequence towards the nation's early child health status.

The dummy for main *source of drinking water* for the household was found to be determining the child health positively *i.e.* those children who live in a household using a safe/protected drinking water have higher health status. Thus, we can argue that a policy that promotes the development of clean water supply in a community also improves the health status of children. Therefore, a government that has a set goal of improving the health status of children with the understanding of its long term consequence on the future generation has to put much effort in improving the accessibility of clean drinking water supply in a community.

Similarly, the dummy for the *type of toilet facility* is found to be the most significant policy variable affecting child health. The result shows that those who use a flush or pit latrine toilet facility have a better health status than otherwise. This result is consistent with the theory that cleaner public environment improves the health status of the people living in the area. It means thus that if a child health state is to be maintained at the highest level the community/household has to have the awareness about the impact of sanitation and hygiene on their health status. This can be achieved by strengthening the capacity of health centers in rural and urban areas in improving the awareness of the community/household about hygiene and sanitation. Thus a development policy that increases the accessibility of health centers and that improves the capacity of existing health center facilities is hence encouraged.

4.2 Determinants of Transitory and Permanent Health.

In this model we try to differentiate between transitory and permanent health status of the child which enabled us to know the impact of the exogenous variables in the short term health state and the long term health state separately. The first latent variable we discussed in the previous section is retained with its indicators and we add the transitory health state by using three other indicators that are a dummy for having *Cough* in the last 24 hours (1 if yes), *a dummy for having Fever* (1 if yes) and dummy for *Vomit everything* in the same time period as before (1 if yes), that are all temporary health states. In this model also the variable *Height-for-Age* has been set equal to one in order to assign a unit of measurement for the latent construct and the same has been done for the dummy variable *fever* in the transitory health state latent construct.[15]

The standardized solutions and t-values are reported below. All the indicators used are significant and are in the expected right direction. The new latent construct transitory health status of children has been well represented by the three indicators *Cough, Fever* and *Vomit* which are all found significant. The basic MIMIC model output estimates and t-values for the two latent variables model is presented in path diagrams in Appendix D.1 and Appendix D.2, respectively. The following Table 4.2 below summarizes the results obtained from the LISREL estimation with the test statistics for evaluating the model fit.

[15] However, in this part of the model the incorporation of additional exogenous variables was constrained by the 15 variables maximum quota set by the student version of LISREL computer package.

Table 4.2: Estimation result of MIMIC model with two latent variables

Variables	C_HEALTH	$\hat{\rho}^{\#}$	CH_TEMPO	$\hat{\rho}^{\#}$
Measurement models			*No. of Observation=1845*	
Weight-for-Age Z-score	0.030*** (4.470)	0.13	-	-
Height-for-Age z-score	1.000[a]	0.40	-	-
Dummy for Comparative Health (1 if same or better)	0.158*** (5.214)	0.52	-	-
Dummy for Fever (1 if yes)	-	-	-1.000	0.36
Dummy for Cough(1 if yes)	-	-	-1.056*** (-19.877)	0.48
Dummy for Vomit (1 if yes)	-	-	-1.060*** (-20.050)	0.40
Structural Equation models				
Dummy for Sex(1 if male)	0.002 (0.193)		-0.003 (-1.291)	
Ratio of Household size to No.of Rooms	-0.008*** (-3.803)		-0.148*** (14.633)	
Dummy for Father's Education (1 if completed primary)	0.441*** (6.554)		0.056*** (7.714)	
Dummy for BCG Vaccination (1 if yes)	0.043** (2.366)		-0.003** (-2.935)	
Dummy for Mother's Education (1 if completed primary)	0.282*** (5.265)		-0.019*** (-3.351)	
Dummy for Decrease in food Availability (1 if yes)	-0.271 (-0.570)		0.049 (1.089)	
Dummy for Source of Drinking water (1 if flush or pit latrine	2.216*** (15.444)		-0.139*** (-2.144)	
Wealth Index	0.232* (1.541)		0.039** (2.544)	
Dummy for Distance from Health center(1 if less than 10km)	0.109 (0.574)		-0.035* (-1.707)	

Goodness of Fit Statistics

Chi-square	677.978 P-value = 0.0000 DF=44
SRMR(Standardized Root Mean Square)	0.048
ECVI(Expected Cross Validation Index)	0.284
AGFI(Adjusted Goodness of Fit)	0.965
RMSEA(Root Mean Squared Error Approximation)	0.081

Note: [a] Normalized to equal one and T-values are presented in parenthesis where *** represents 99% significance level, ** 95% and * 90%
refers to standardized multiple correlations.

In the above model the squared multiple correlation ($\hat{\rho}$) for the permanent health status remained the same as before while for the transitory health status $\hat{\rho}$ is relatively better. The highest $\hat{\rho}$ is 0.48 for the dummy variable *cough* which implies that 48 percent of the variation in having fever is explained by the latent transitory health status. We have 0.40 and 0.36 for the variables *vomit* and *fever*, respectively.

The results show that the overall fit of the model is good as indicated by Adjusted Goodness of Fit Index (AGFI), Root Mean Squared Residual (RMR), and the significance of chi square. A very high AGFI, significant chi square and low RMR and ECVI suggest that models fit the data very well. The signs in the measurement model are in the expected right direction and their effects are significant. The Table 4.2 above shows all the results of the basic model estimates in their standardized solutions.

In the structural equations, except *Decrease in Food Availability* and *Healthcare Access* all the rest of the variables appear to be significant factors that affect the child's permanent health state between 95 and 99 percent level of significance. On the other hand, except *Decrease in Food Availability* and *Sex* all the rest of the variables appear to be significant factors affecting transitory health state with between 95 and 99 percent level of significance.

Before proceeding with the interpretations, an inquiry about the existence of a plausible adjustment in the model that would increase the fit of the model was made using the modification indices that is available in LISREL. The modification index is calculated by completely changing the indicators used for one latent variable to the other. In this paper's particular case the indicators for the permanent child health were taken to represent the transitory child health and the indicators of the transitory child health to the permanent one. The modification indices for the above MIMIC model are presented in the path diagram in Appendix D.3. The Table 4.3 below also shows the result of the modification indices.

Table 4.3: Modification indices for the two latent variable MIMIC models

Variables	C_HEALTH	CH_TEMPO
Weight-for-Age z-score	- -	2.398
Height-for-Age z-score	- -	53.510
Comparative Health	- -	227.433
Dummy for Fever (1 if yes)	1.110	- -
Dummy for Cough 91 if yes)	5.573	- -
Dummy for Vomit (1 if yes)	2.320	- -

The Table 4.3 shows the index of change that would come if we exchange the indicators used for the two latent variables. The largest modification index is for the variable *Comparative Health* which is equal to 227.433. This implies that it gives a significant better fit to the model if *Comparative Health* is used as an indicator of child transitory health status than the permanent health status. This suggestion is acceptable and may have come from the difference in the unit of measurement of the variable *Comparative Health* with those of *Height-for-Age* and *Weight-for-Age z-scores.*

Accordingly, the variable *comparative health* has been moved to indicate the latent variable transitory child health. This has actually improved the fit of the model significantly as presented below. The newly estimated result is presented in a path diagram in Appendix D.4 with the t-values in Appendix D.5 and the summary table is presented below;

Table 4.4: Estimation result of MIMIC model with two latent variables after modification

Variables	C_HEALTH	$\hat{\rho}^{\#}$	CH_TEMPO	$\hat{\rho}^{\#}$
Measurement model			*No. of Observations= 1845*	
Weight-for-Age z-score	0.031*** (4.38)	0.13	- -	
Height-for-Age z-score	1.000ⁿ	0.38	- -	
Dummy for Comparative Health(1 if same or better)	-		3.253*** (15.527)	0.17
Dummy for FEVER (1 if Yes)	-		-1.000ⁿ	0.36
Dummy for COUGH (1 if Yes)	-		-1.020*** (-21.089)	0.46
Dummy for VOMIT (1 if Yes)	-		-1.028*** (-21.140)	0.40
Structural model				
Dummy for Sex (1 if Male)	0.007 (0.012)		-0.003* (1.639)	
Ratio of Household size to No. of Rooms in the household	-0.006 (0.079)		-0.158*** (-15.869)	
Dummy for Father's Education (1 if completed primary)	0.391*** (5.485)		0.058*** (7.962)	
Dummy for Mother's Education (1 if completed primary)	0.344*** (6.124)		-0.021** (-3.675)	
Dummy for BCG Vaccination (1 if yes)	0.032* (1.813)		0.004** (3.355)	
Dummy for Decrease in Food Availability (1 if yes)	0.140 (0.283)		-0.032 (-0.725)	
Dummy for Source of Drinking Water (1 if safe/protected)	2.292*** (15.707)		0.025* (1.824)	
Wealth Index	0.075 (0.521)		0.030** (2.092)	
Dummy for Healthcare Access (1 if < 10km)	0.085 (0.432)		0.015 (0.756)	
Test statistics				
χ^2 (Chi-Square)	428.430		DF=44 P=0.0000	
ECVI(Expected Cross Validation Index)	0.204			
RMSEA(Root Mean Squared Error of Approximation)	0.0632			
SRMR(Standardized Root Mean Square)	0.0232			
AGFI(Adjusted Goodness of Fit Index)	0.997			

Note: ⁿ Normalized to equal one and T-values are presented in parenthesis where *** represents ~~99%~~ significance level, ** ~~95%~~ and * ~~90%~~
[#] refers to the squared multiple correlations.
1 5 10

As the Table 4.3 above shows the estimates of the test statistics representing the fit of the model has greatly improved when the variable *comparative health* is used as the indicator of the latent transitory health status. This has been reflected in the improvement of the AGFI from 96.5 to 99.7 percent and the fall of the value of RMSEA, SRMR and ECVI which shows the improvements in model fit.

In this model also, the measurement model's indicator variables are in the expected right direction and they are significant at 99 percent significance level. The interpretation of the above model is organized as follows.

Household/Individual characteristics

The significance of the variable *Sex* shows that being male has a negative relation with transitory and though insignificant a positive relation with permanent health state. This is may be true male children are more exposed to the outer environment and therefore, may become more prone to temporary illness. Shehzad (2006) argued that there are no conclusive results for the direction of gender bias and it is clear that girls are not as often taken to the health care facilities as compared to boys and receive less attention from their parents. The result of this thesis also confirm that although, boys are more prone to temporary health illness due to more outer exposure, they may have accessed better health care facility as compared to girls which justifies the reduction in the significance of the factor loading for the permanent health state.

In this estimation result interpretation, however emphasis has been placed to the impact of *Sex* on the transitory health state rather than on the permanent health state as the indicators in the permanent health state are age and gender standardized.

The *ratio of household size to the total number of rooms in the household* shows that the higher the number of individuals lives in one room the lower the health status of children in both their permanent and transitory states but it is significant only in transitory health status. The effect of household size per room on child's both permanent and transitory health status being negative is may be a result of congestion effects and lower income share per head in the household. Our result is consistent with the findings of Christiaensen and Alderman (2001), Wolfe and Behrman (1981), Tirfe (2006) and others. This result proves the fact that transitory health status is determined by temporary situations/events in the household.

Parental Education

In this model both father's education and mother's education are found to be significant factors affecting the permanent health state positively. That means those children born from educated parents will have a better permanent health. We interpret these results to reflect the efficiency of the women in household production, with the above mentioned caveat about schooling representing tastes and genetics. However, in this model the transitory health is affected negatively by the mother's education. This can be because the educated mother has some job to do outside the household and has less time to follow up her child for temporary health deviations.

On the impact of mother's education on the permanent health status, it can also be because more educated parent produce better child health as a result of their improved awareness of health related matters. They are also more capable of using the medical advice when children are sick and can help in quick recovery by taking care of nutrition and food intake of children. Similarly, more educated mothers are better producers of child health. The effect of education works through increased awareness and scientific understanding of disease causation. Better-educated mothers appear to demand fewer high quality (healthy) children as compared to more unhealthy children.

Household wealth

The wealth index that shows the household economic status is significant only in determining the child's temporary health state. This may have come from the fact that those households with a higher economic status have the ability to pay for medical care costs for their children to recover from temporary illnesses. On the other hand, children born from a household with a lower economic status have to wait until their illness gets worse and may not even visit a medical center what so ever.

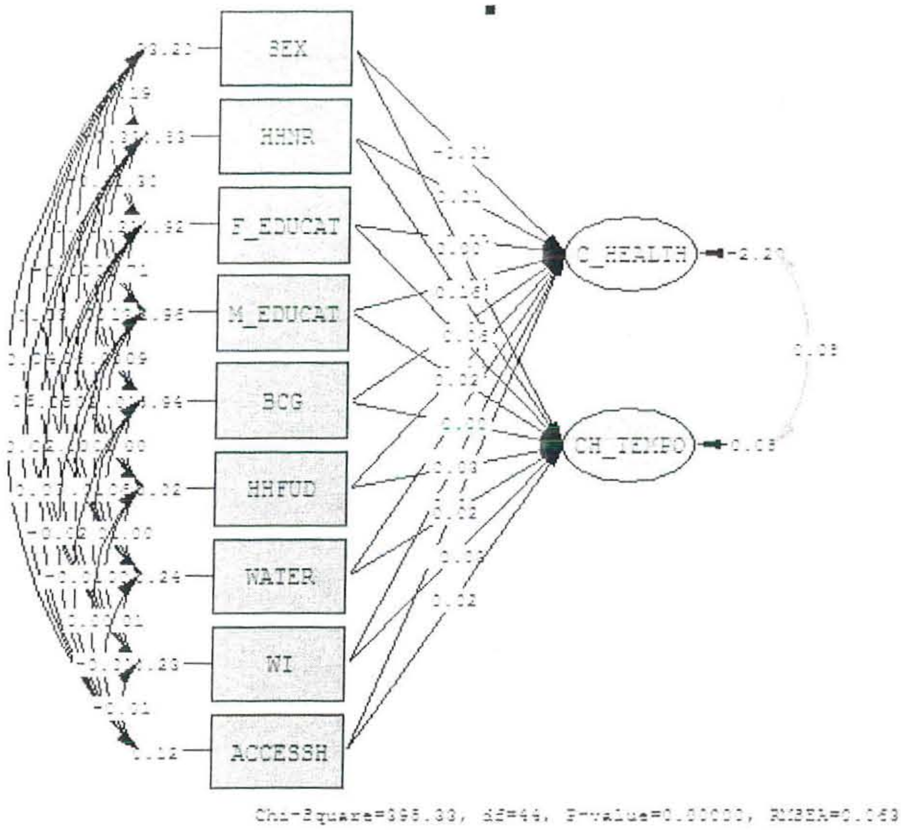
Government policy variables/Public Environment

Here in this model the BCG vaccination has been taken in place of Measles Vaccination as it became significant. The model's output depicted in Table 4.4 shows that BCG vaccination is a significant factor determining child health's transitory health at 95 percent significance level and permanent health at the 90 percent significance level. The significance of the variable BCG vaccination shows the importance of expanding the availability of health center providing this service in the community. The vaccination variable is more significant in affecting a child's transitory health status than permanent health as it helps them resist diseases. Taking it as the indicator of the impact of policies on children, one can argue that government policy that improves the level of health service delivery in the community improves the health status of children.

A similar output is obtained for the *source of drinking water.* However, the result obtained for the source of drinking water shows that it is much more significant in determining the children' permanent health than their transitory health. This is may be because the lack of pure drinking water leads to temporary health deviation but this might have led to a cumulated disorder in the permanent health state of the child.

The model's output also showed that child's permanent health state has a positive and significant effect on child's transitory health as seen from the standardized correlation between the two latent variables in the structural model's path diagram presented below. This means that children with better permanent health can cope with transitory health deviations.

Figure 4.1: Standardized correlation between the two latent variables in the structural model.



The above figure shows that there exists a very high correlation between the permanent health status and the transitory health status. The t-statistics (not reported) also shows that the correlation is significant at 1 percent level. This implies that public policy makers should target those determinants that significantly affect the permanent health status as this would help to indirectly affect the transitory health state. In other words, policies or interventions that positively affect the permanent health state can significantly improve also the transitory health state since those with strong permanent health state will be able to cope with temporary health deviations and the vice versa. It also means that a good permanent health status enables the children to overcome morbidity and helps in quick recovery.

Chapter V

Conclusion and Policy Implication

5.1 Summary and Conclusion

The correlates of early childhood health status were identified by using a simultaneous equation analysis of the multiple causes and multiple indicators model.

In this particular paper an estimate was provided with a mixed demand-production function for child health in Ethiopia. This is because child health is assumed to be determined by some inputs from the structural production function and some variables from the reduced form demand relations.

The results are comparable to other studies and also add to the understanding of the correlates of child health in Ethiopia by using new techniques. Structural equation models have been estimated with unobservable variables. Such techniques have been successfully used among others by Wagstaff (1986, 1993), Van Doorslaer (1987), Behrman and Wolfe (1987), Kiinshnin (2003), Shehzad (2004, 2006) and Andreas and Freidrich (2008).

Joreskog and Sorbom (1989) show that LISREL uses confirmatory factor analysis and structural equations typically associated with econometrics to explain latent (unobservable or theoretical) variables. A special case of these relations is MIMIC models that links the two parts of structural relations estimated in LISREL; the measurement part and the structural part. The estimation of MIMIC models for child health enables to capture complex relations in the measurement model and see how their validity is affected after the introduction of the structural equation model. The method of estimation is weighted least squares that provides consistent estimates of parameters under the assumption of normality and is quite robust against departures from normality. Results of MIMIC models show that unobservable child health states can be represented by several observable health indicators. The measurement model uncovers complex and diverse relationships that exist between observable health indicators but unobservable health states.

Other studies that estimate child health for example, Thomas *et al.*, (1991), Barrera (1990), Alderman and Garcia (1994) use a single health indicator as height or weight for age as the dependent variable. However, as Behrman and Deolalkar (1988, p. 650) observe, "*the measures differ significantly in regard to their costs and measurement errors and may represent different dimensions of health status*". Hence, they urge the need to use multiple indicators to represent health status in an empirical analysis. Outputs of this paper show that the use of MIMIC models lead to a more comprehensive understanding of the determinants of child health as compared to other studies that rely on single health measures.

The Young Lives data contains rich information on relevant child health indicators and results identify significant effects of family size, housing characteristics on transitory health distortions. Child illnesses such as cough and fever are affected by *family size to number of rooms*, housing and parental education. Results imply that incidence of these illnesses can be reduced by improving housing characteristics (piped water, sanitation etc). Child's permanent health has a significant positive impact on child's transitory health deviations. A good permanent health status enables the children to overcome morbidity and helps in quick recovery. However, factors like family size can produce more respiratory infections through increased *person-to-person* contact. The results support the view that children from large families have deteriorated permanent health and the effect operates through lower per head income.

The result of this paper supports Grossman's technical efficiency hypothesis: *education acts as an efficiency increasing factor in producing health.* This is because more educated parents are able to produce more health output for a given level of use of health inputs in both of the models we estimated. This is because the argument that more educated parent, particularly that of the mother, produce better child health because of their improved awareness of health related matters has been confirmed in this study.

Similarly, more educated mothers are better producers of child health. The effect of education works through increased awareness and scientific understanding of disease causation. The significant effects of these variables are evident but sometimes, may result in collinear effects. In structural equation models, a very high multi-collinearity is unlikely to happen and is assumed absent as this will result in singular covariance matrices that cannot perform certain calculations (Matrix inversion) and prevent structural equation estimation. However, in present result of this paper, multi-collinearity has not been found to be a significant problem.

This thesis work found that the *ratio household size to the total number of rooms* in the household a negative relation with the health status of children in both their permanent and transitory states. This can be a result of congestion effects and lower income share per head in the household.

The *wealth index* that shows the household economic status is significant only in determining the child's temporary health state. The significance of the variable *BCG vaccination* shows the importance of expanding the availability of health center that provide the vaccination in the community. The vaccination variable is more significant in affecting a child's transitory health status as it helps them resist diseases.

The model's output also showed that child's permanent health state has a positive and significant effect on child's transitory health as seen from the standardized correlation between the two latent variables in the structural.

Finally, from the evidences given one can conclude that, despite the existing constraints, *parental education, household size, source of drinking water, vaccination availability and accessibility* etc are among the important correlates of early childhood health outcomes in Ethiopia. These results have implications for social policy, where improvement in the above mentioned variables is positively linked to improvement in early childhood health and better overall health state in the future.

5.2 Policy Implication

The findings of the empirical analysis show that among others parental education, infrastructural development in terms of health facility and household size are the most important determinants. The household food availability and the economic status of the household are also significant determinants at different levels of significance. The study also confirmed that the causes of poor health in childhood are complex, multi-dimensional and interrelated requiring a careful understanding of the problem to implement effective policies and programs.

Consequently, intervention programs on parental education or in general parent's status would contribute to the effort towards improving the health status of children. In particular, considering the low level of mother's education at national level and especially in rural parts of the country, policy actions that are meant to improve the educational status of women are critical in addressing the problem through improving their income earning capacity and also enhancing the quality of care and attention they provide to their children. From theoretical argument perspective, with a caveat about identifying efficiency from generic or tastes effects, we believe that women's schooling represents an important mechanism for improving child health through increasing efficiency in household production. But for such schooling the investment period is too long (Wolfe and Behrman, 1981). Adult education programs directed towards health practices may be more efficient, although we don't have data to test this possibility.

In addition to that, government policy makers are advised to design and implement a population policy which controls for large family size as the study found that the ratio of household size to the number of rooms is a significant correlate of child health.

Our result shows that good quality basic services and in particular primary health and drinking water are an essential complement to economic strategies aiming to promote child health. Given the critical importance and life long effects of good health they play an especially critical role for children. In most contexts, even with a decline in state expenditure on basic services due to austerity programmes or war, the state continues to have primary responsibility for co-ordinating and providing such services and may well be the only organisation able to support and maintain services, particularly in isolated areas. Support for basic social services, should, therefore concentrate on supporting state capacity. Where the state is unable to provide basic social services for all, alternatives - provided by NGOs, religious organisations, or grassroots organisations are essential. Whilst in some contexts, such organisations play a vital role in providing services in under-served areas as among private and church schools that may be increasing inequities by providing a superior service to those who can afford to pay.

Aid can play an important role in assisting governments lacking financial capacity to deliver good quality basic services to all families and children. Debt relief linked to expenditure on social services would clearly be of immense benefit to children. A reduction in military expenditure - which in many countries will depend on conflict resolution - and investment of the resources freed up in basic services is a related priority.

Within sectoral budgets, the distribution of spending is of profound significance for improving child health. To benefit the majority of poor children, expenditures must be oriented to primary health care, safe rural water supply and sanitation.

These insights, therefore with the support of other studies provide a better basis for prediction and policy analysis regarding child health in Ethiopia- and thus for improving current and long-run welfare.

Bibliography

Alderman, H. (1990). '*Nutritional status in Ghana and its determinants*', World Bank, Washington DC., USA

Alderman, H. and M. Garcia (1994), '*Food Security and health security: explaining the levels of nutritional status in Pakistan*', Economic Development and Cultural Change 42(3) 485-507

Alemu M., Jones N., Bekele T. (2005a) '*Tackling Malnutrition In Ethiopia: Do Sustainable Development Poverty Reduction Programme's Underlying Assumption Reflect Local Realities?*' Working Paper No. 19, Young Lives, Save the Children UK.

Alemu M., Bekele T, Tassew W., Nicola J., John S., Tekie A., Getachew A (2005b), *Child Nutritional status in poor Ethiopian households: The role of gender, assets and location* Working paper No. 26, Young Lives, Save the Children UK.

Almond, D. (2005)*; Is the 1918 influenza pandemic over? Longterm effects of in utea influenza exposure in the post 1940 US population*, Journal of political economy, 672-712.

Andreas Buhn and Freidrich Shnieder, (2008), '*MIMIC Models, Co-Integration and Error Corrections: An Application to the French shadow economy.*' Journal of Economic Literature 38/1, pp. 68-101.

Attanasio, O., Gomez, L.C., Hernandez, M.V. (2005) '*Child Health in Rural Colombia: Determinants and Policy Interventions*' Centre for the evaluation of Development Studies, The Institute for Fiscal Studies. Report summary: Familias 03.

Avery, R. B. (1979), '*Modeling monetary policy as an unobserved variable'*, Journal of Econometrics, 10

Ayalew D. (2006), '*Determinants of child Nutritional status in Ethiopia: a longitudinal analysis'*, M.sc Thesis, Department of Economics, Addis Ababa University.

Barrera, A.: (1990), '*The role of maternal schooling and its interaction with public health programs*', Journal of Development Economics 32(1), pp. 69–92.

Behrman J. and Rosenzweig M., (2002) '*Does increasing mother's schooling raise the schooling of the next generation?*' American Economic Review, 323-334

________________________(2004), '*Returns to birth weight*'. Review of economics and statistics 586-601

Behrman, J. (1990). '*Women's Schooling and Non-Market Productivity: A Survey and Reappraisal.*' Washington, DC, United States: World Bank. Mimeographed document.

Behrman, J. R. and A. B. Deolalikar: (1988), *Health and nutrition,* in H. Chenery, and T. N. Srinivasan (eds.), Handbook of Development Economics (North Holland, Amsterdam).

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): 185-204.

Black S. Devereux, P.Salvanes, K. (2005), *Why the apple does not fall far: understanding intergenerational transmission of human capital.* American Economic Review, 437-449

Blackman, R. and Litchfield, J. (2001) '*Child Poverty*' Poverty Research Unit at Sussex,No.1

Bollen, K. A. (1989). '*Structural Equations with Latent Variables*'. Wiley, New York, first edition. Accessed from a web based preview.

__________ (2002). "*Latent Variables In Psychology And The Social Sciences.*" Annual Review of Psychology 53(1): 605-634.

Boniface, D. R. and Tefft, M. E. (1997) '*The application of structural equation modeling to the construction of an index for the measurement of health related behaviors*'. Statistician, 46(4):505–514.

Case A. Lubotskey, D. Paxson (2002) '*Economic Status and health in childhood: the origins of the guardian*' American Economic Review 1308-1334

Case A., Fertig A, Paxton, (2005) '*The lasting impact of childhood health and circumstance*', Journal of health economics. 365-389.

Chaudhuri, A. (2003) '*Program impact on the Health and Nutritional Status of Children: Evidence from Bangladesh*' University of Washington, Department of Economics, doctoral thesis..

Christian Helmers and Manasa Patnam '*The Formation and Evolution of Childhood Skill Acquisition: Evidence from India.*' Young Lives project Working Paper No. 38, Department of International Development, Oxford University.

Christiaensen, L. and Alderman, H. (2001) '*Child malnutrition in Ethiopia: Can Maternal knowledge Augment the Role of Income*' Africa Region working paper series, No 22.

Currie and Hyson, (1999) '*Is the impact of health shocks cushioned by socioeconomic status?The case of low birth weight.*' American Economic Review, Papers and proceedings 245-250

Currie and Madrian, (1999) *Health, Health insurance and the labor market*, In: Orley Ashenflter, David Card (Eds.), In: Handbook of Labour Economics, Vol 3A, Amsterdam, North-Holland.

Currie and Moretti, (2003) '*Mother's education and the intergenerational transmission of human capital: evidence from college openings*'. Quartely Journal of Economics..

Currie and Stablie, 2003. '*Socioeconomic Status and child health: why is the relationship stronger for older childen*?' American Economic Review, 1813-1823.,

David Card (eds), In *Handbook of labor Economics*, Vol 3.A Amsterdam, North-Holland;

Dercon, Stefen and Hoddinot, John (2003) ' *Health, Shocks and poverty Persistence in Ethiopia'* World Institute of Development Economics Research, Helsinki, part of the United Nations University.

Desai and Alva, (1999) '*Maternal Education and child health: Is there a strong causal relationship?* Demography 71-81;

Dwyer D, Bruce J. (1988) *A Home Divided, Women and Income in the Third World.* Stanford University Press, Stanford CA. Epidemiology and Community Health 40.

Ersland M, Ried W, Ulrich V. (1995) '*Health, health care, and the environment. Econometric evidence from German micro data.*' Health Economics 15: 169-182.

Escobal, J. Saavedra, J.,Suaez, P., Huttly, S.,Penny, M. lanata,C., Villar, E. (2005) '*The Interaction of Public Assets, private Assets and Community characteristics and its effects on early childhood Height for age in Peru*' working paper no. 14, Young Lives, Save the Children, UK

Fan, X., Thompson, B., and Wang, L. (1999) '*Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes'*. Journal of Structural Equation Modeling, 6(1):56–83.

Gertler, P. and J. Gruber (2002). "*Insuring consumption against illness.*" American Economic Review: 51-70.

Giuffrida A., R.F. Iunes and W.D. Savedoff (2005), '*Health and poverty in Brazil: Estimation by structural equation model with latent variables*', Technical note on health No.1, Inter-American development Bank, Washington DC, USA

Glewwe, (1999) '*Why does mother's schooling raise child health in developing countries? Evidence from Moroco.*' Journal of Human resources 10(1)124-159

Glewwe, Jacoby and king, (2001), '*The impact of income growth and provision of health care services on child nutrition in Vietnam'*, Journal of Human Resources. 23(2) 212-

Greene (2003). *Econometric Analysis*. Macmillan, New York, third edition.

Grossman, M. (1972a). '*The Demand for Health: A Theoretical and Empirical Investigation*' Number 119 in Occasional Paper. National Bureau of Economic Research, New York.

Grossman, M. (1972b). '*On the concept of health capital and the demand for health'*. Journal of Political Economy, 80(2):223–255.

Griliches, Z. and W. M. Mason (1972). "*Education, income, and ability.*" The Journal of Political Economy: 74-103.

Hakkinen, U. (1991). '*The production of health and the demand for health –care in Finland.*'Social Science and Medicine, 33(3):225–237.

Hancock GR, Nevitt J. (1999) *'Bootstrapping and the identification of exogenous latent variables within structural equation models'. Structural Equation Modeling health, nutrition, medical care usage, and household sanitation*?', Journal of Econometrics.

Harper A., Bhat, N., J. G. Wright, (2003). *Influenza-associated deaths among children in the United States*, 2003-2004. 353: 2559-2567

Harpham T. (2003)., '*Measuring child poverty and health: a new international study*', Young Lives working paper No. 2, London.

Harper C. and Marcus R. (1999/00). '*Childhood poverty in sub Saharan Africa'*, The Save the Children Fund(UK), London.

Harvey, A.C. (1981): '*Time series Models'*. (Halsted press, New York)

Hauser (1972), '*The treatment of unobservable variables in path analysis'*, in sociological methodology,

Jensen T. and Richter, K. (2001). "*Understanding the relationship between poverty and children's health*" European Economic Review 45(4-6): 1031-1039

Joreskog, K. G: (1971), '*Statistical analysis of sets of congeneric tests*', Psychometrica

Jöreskog KG, Sörbom D. (1984). *LISREL-IV User's Guide*. Scientific Software, Mooresville, Indiana.

Joreskog, K. G. and D. Sorbom: (1989), *LISREL 7. A Guide to the Program and Applications*. 2nd edn. (SPSS Inc.).

Joreskog, K. G. and S¨orbom, D. (1993a). '*LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*.' Scientific Software International, Inc., Chicago, IL, first edition.

Joreskog, K. G. and S¨orbom, D. (1993b). *LISREL 8 User's Reference Guide*. Scientific Software International, Inc., Chicago, IL.

Joreskog, K. G. and S¨orbom, D. (1993c). *PRELIS 2 User's Reference Guide*. Scientific Software International, Inc., Chicago, IL.

Joreskog, K. G. (1994). '*On the estimation of polychoric correlations and their asymptotic covariance matrix*'. Psychometrika, 59(3):381–389.

Joreskog, K. G. (2001). '*Analysis of Ordinal Variables: Longitudinal Data*'. Scientific Software International, Inc., Chicago, IL.

Kurz, K. M, and C. Johnson-Welch: (1997), ''*Gender differences in the health and nutrition of under-fives show prevalent bias against girls*''. Basic Support for Institutionalizing Child Survival (BASICS II), USAID

Kimhi A. (2002)., '*Socioeconomic Determinants of health and physical fitness in southern Ethiopia*'., The Center for Agricultural Economics Research, Discussion No. 5. 02, the Hebrew University.

Kiiskinen, U., (2003) '*A Health Production Approach to the Economic Analysis of Health Promotion*.' PhD Dissertation, Department of Health and Functional Capacity National Public Health Institute, Helsinki, Finland and Department of Economics and Related Studies, University of York, United Kingdom.

Lahiri, K (1976) '*Inflationary expectations: their formation and interest rate effects*' American Economic Review 120.

Laurence HL. (2008), '*A Multiple Indicators and Multiple Causes (MIMIC) Model of Immigrant settlement success*,' Working paper No. 160, National Institute of Labour studies, Flinders University, Adelaide, Australia.

Manning WG, Newhouse JP, Ware JE Jr. (1982) '*The Status of Health in Demand Estimation: Beyond Excellent, Good, Fair and Poor*'. In Fuchs V. (eds) Economic Aspects of Health. University of Chicago Press, Chicago.

Maxwell, D., Levin, C., Kelmesu,M.A., Reul,M., Moris,S. and Ahiadeke, C. (2000), '*Urban Livelihoods and food and nutrition security in Greater Accra, Ghana*' International Food policy Research Institute, research Report No. 112.

Mehrotra, S. (1998). "*Education for all: Policy Lessons from high Achieving countries.*" International Review of Education 44(5):461-484

Mullahy J. (1999) '*It'll only hurt a second? Microeconomic determinants of who gets flu shots'*. Health Economics, Amsterdam: North-Holland.

Ministry of Finance and Economic Development, (2002a) *Sustainable Development and Poverty Reduction Programme, Ethiopia,* Addis Ababa, Ethiopia.

Ministry of Finance and Economic Development, (2002b), *Development an Poverty Profile of Ethiopia,* Welfare Monitoring Unit, Addis Ababa, Ethiopia.Addis Ababa, Ethiopia

Perlman M. (eds.) *Health, Economics, and Health Economics*. North-Holland, Amsterdam.

Sahn, D.E and Stifel, D.C, (2002) 'Parental Preferences for Nutrition of Boys and girls: Evidence from Africa' the Journal of development Studies, Vol. 39, No.1

Robinson P. and M.C Ferrara (1977) '*The estimation of a model for an observable variable with endogenous causes.*', Publisher not specified.

Rosenzweig, M. R. and T. P. Schultz (1983). "*Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight.*" The Journal of Political Economy: 723-746.

Schultz, T. P. (2002). "*Wage gains associated with height as a form of health human capital.*" American Economic Review. 349-353

__________ (eds.) *Wealth from Health. Linking Social Investments to Earnings in Latin America.* Inter-American Development Bank, Washington DC.

Schultz TP, and Tansel A. (1997). '*Wage and labor supply effects of illness in Cote d'Ivoire and Ghana: instrumental variable estimates for days disabled'*. Journal of Development Economics 53(2).

Senauer and Garcia (1991), "*Determinants of the nutrition and health status of pre-school children: an analysis with longitudinal data*" Economic Development andciltural change. Vol. 39 No. 2371-389.

Setboonsarnge, S. (2005). "*Child Malnutrition as a poverty indicator: an evaluation in the context of different development interventions in Indonesia.*" Asian Development Bank (ADB) Institute Discussion paper 21.

Shehzad S.: (2004), '*A Measurement Model of Child Health: Latent variables approach,* Journal of Health and Population in the Developing Countries'. 85(8) 253-280.

_________(2006) 'The *determinants of child health in Pakistan: An economic Analysis'*, Social Indicators Research 78(3) 531-556.

Sorbom, D and K. G. Joreskog: (1981), '*The use of LISREL in sociological model building'*, in E. Borgatta and D. J. Jackson (eds.), Factors Analysis and Measurement in Sociological Research: A Multidimensional Perspective (Beverly Hill, Sage)

Strauss,J. (1990)'*Households, Communities and preschool child nutrition outcomes: evidence from cote d'Ivioure*'. Economic Development and cultural change, 10(1) 231-261;

Strauss J., D. Thomas *et al.* (1994). '*Human Resources: Empirical modeling of household and family decision,*' RAND.

Strauss, J. and D.Thomas (1998), "*Health, Nutrition and Economic Development.*" Journal of economic Literature

Tassew W,. Alemu. M, et al. (2008). *Young Lives Ethiopia: Round 2 Survey Report*, Oxford: Young Lives.

Tekie Alemu, Getachew Asgedom, Janet Liebenberg, Alemu Mekonnen, John Seager, Bekele Tefera and Tassew Woldehanna, (2003) Young Lives Preliminary Country Report: Ethiopia

Tanaka, J. S. (1993). *'Multifaceted conceptions of fit in structural equation models.'* In Bollen, K. A. and Long, J. S., editors, Testing Structural Equation Models, Sage Focus Edictions. Sage, Newbury Park.

Thomas, D. Strauss, J. Henriques,.(1990) *'Child Survival, height for age and household characteristics in Brazil.'* Journal of Economic Development

Thomas, D., J. Strauss and M.M Henriques. (1990). "*Child Survival, Height for Age and Household Characteristics in Brazil.*" Journal of Development Economics 33: 197-234.

Thomas, D., Strauss, J., Henriques, M.-H., (1991) '*How does mother's education affect child height*? 'Journal of Human resource, 183-211.;

Tirfe, M. (2006).' *The Role of Maternal Characteristics on Nutritional Status of Ethiopian Children'*, Addis Ababa University, unpublished Msc Thesis.

Todaro, M. and Smith, S. (2005) '*Economic Development*', ninth Edition. Pearson-Addison Wesley, Boston. Tenenhaus.

UNICEF (2005)' *Child Poverty in rich Counties'* Innocent Research center, New York NY.

________(1998). "*The state of the world's children*" New York:.

Van de Ven,. and Van der Gaag, J. (1982). *'Health as an unobservable: A mimic model of the demand for health care.'* Journal of Health Economics, 1(2):157–183.

Van Der Gaag and Wolfe (1981) '*Estimating demand for medical care- health as a critical factor for adults and children,*' Econometrics of health care

Van der Gaag, J. and Wolfe, B. (1991). '*Estimating demand for medical care - health as a critical factor for adults and children.*' In Duru, G., editor, Econometrics of health care, pages 31–58. Kluwer Academic Publishers, Dordrecht.

Van Doorslaer, E. K. (1987). '*Health, Knowledge and the Demand for Medical Care: an econometric analysis.*' Van Gorcum, Assen, The Netherlands.

Van Vliet, R. C. J. A. and Van Praag, B. M. S. (1987). '*Health status estimation on the basis of mimic-health care models.*' Journal of Health Economics, 6(1):27–42

Wagstaff A. (1986). '*The demand for health: Some new empirical evidence.*' *Journal of Health Economics* 5:

Wagstaff, A.: (1986), '*The demand for health: Theory and applications*', Journal of Epidemiology and Community Health 40, pp. 1–11.

Wagstaff, A.: (1993), '*The Demand for Health: An empirical reformulation of the Grossman model*', Econometrics and Health Economics 2, pp. 189–198.

Weeks, (2001), *Models of Unobserved Latent variables*, Advanced Econometrics, Faculty of Economics and Politics, University of Cambridge, Notes

WHO (1985). '*Targets for Health for All.*' World Health Organization, Regional Office for Europe, Copenhagen.

WMS, (2004) '*Welfare Monitoring Survey: Analytical report.*' Statistical Bulletin 339-A. The Federal Democratic Republic of Ethiopia, Central Statistics Authority, Addis Ababa.

Woldemariam Girma and Timotious Genebo (2002)'*Determinants of the Nutritional Status of Mothers and Children in Ethiopia*' Calverton, Maryland, USA: ORD Macro.

Wolfe BL, Behrman JR. (1984).' *Determinants of women's health-status and health-care utilization in a developing-country: a latent variable approach.* Review of Economics and Statistics 66(4):

Wolfe BL, van der Gaag J. (1981). *A New Health Status Index for Children.* In van der Gaag J, Perlman M. (eds.) *Health, Economics, and Health Economics*. North-Holland, Amsterdam

World Bank (1999) '*Voices of the Poor and Voices of Children and Adolescents in Latin America and the Caribbean,*' Washington DC, USA.

________ (2000/2001) *World Development Report: Attacking Poverty*, Washington DC, USA.

________ (2006) '*Repositioning Nutrition as central to Development- A strategy for Large Scale Action*' The International Bank for Reconstruction and Development/ The World Bank, Washington DC, USA.

World Health Organization (2003). *World Health Report*: Shaping the Future, Geneva, Swizerland.

Yuyu Chen and Hongbin Li, (2009); *Mother's education and Child Health: Is there a nurturing effect?* Journal of Health Economics 413-426

Zellner (1970), '*Estimation of Regression Relationships Containing Unobservable Variables'* International economic review, 11

Appendices

Appendix A: Deriving the MIMIC model's covariance Matrix

The MIMIC model's structural and measurement equations are $\eta_i = \gamma' x_i + \varsigma_i$ and $y_i = \lambda\eta_i + \varepsilon_i$, respectively. Expressing this model in terms of covariances gives:

$$\Sigma = \begin{pmatrix} Var(y_i) & Cov(y_i, x_i) \\ Cov(x_i, y_i) & Var(x_i) \end{pmatrix} = \left(\begin{bmatrix} y_i \\ x_i \end{bmatrix} \begin{bmatrix} y_i \\ x_i \end{bmatrix}' \right).$$

After taking the transposes, multiplications, and making use of the assumptions that:

1. The variances are measured as deviations from mean, i.e.

   $E(\eta_i) = E(x_i) = E(\varsigma_i) = E(y_i) = E(\varepsilon_i) = 0$

2. The error terms do not correlate to the causes, i.e. $E(x_i\varsigma'_i) = E(\varsigma_i x'_i) = 0$ and

   $E(x_i\varepsilon'_i) = E(\varepsilon_i x'_i) = 0;$

3. The error terms do not correlate across equations, $E(\varepsilon_i\varsigma'_i) = E(\varsigma_i\varepsilon'_i) = 0;$ and

4. The errors of the measurement model do not correlate to the latent variable,

   i.e.

   $E(\eta_i\varepsilon'_i) = E(\varepsilon_i\eta'_i) = 0;$

We distribute the expectation operator and can thus derive both the variance and covariance between the observable variables. By doing this, it follows that:

$$\begin{aligned} E(y_i y'_i) &= E\left[(\lambda\eta_i + \varepsilon_i)(\lambda\eta_i + \varepsilon_i)'\right] \\ &= E(\lambda\eta_i\eta'_i\lambda' + \lambda\eta_i\varepsilon'_i + \varepsilon_i\eta'_i\lambda' + \varepsilon_i\varepsilon'_i) \\ &= \lambda E(\eta_i\eta'_i)\lambda' + \Theta_\varepsilon \\ &= \lambda E\left[(\gamma' x_i + \varsigma_i)(\gamma' x_i + \varsigma_i)'\right]\lambda' + \Theta_\varepsilon \\ &= \lambda E(\gamma' x_i x'_i\gamma + \gamma' x_i\varsigma'_i + \varsigma_i x'_i\gamma + \varsigma_i\varsigma'_i)\lambda' + \Theta_\varepsilon \\ &= \lambda(\gamma'\Phi\gamma + \Psi)\lambda' + \Theta_\varepsilon \end{aligned}$$

$$
\begin{aligned}
E(x_i y'_i) &= E\left[x_i(\lambda\eta_i + \varepsilon_i)'\right] \\
&= E(x_i\eta'_i\lambda' + x_i\varepsilon'_i + \varepsilon_i\eta'_i\lambda' + \varepsilon_i\varepsilon'_i) \\
&= E(x_i\eta'_i)\lambda'
\end{aligned}
$$

$$
\begin{aligned}
&= E\left[(x_i)(\gamma' x_i + \varsigma_i)'\right]\lambda' \\
&= E(x_i x'_i\gamma + x_i\varsigma'_i)\lambda' \\
&= \Phi\gamma\lambda' \\
E(y_i x'_i) &= (\Phi\gamma\lambda')' \\
&= \lambda\gamma'\Phi \\
E(x_i x'_i) &= \Phi
\end{aligned}
$$

Thus, Θ_ε is the covariance matrix of the error terms in the measurement model; Ψ is the variance of the error term in the structural equation; and, Φ is the covariance matrix of the causes. Finally, the covariance matrix of the MIMIC model is:

$$
\Sigma = \begin{pmatrix} \lambda(\gamma'\Phi\gamma + \Psi) + \Theta_\varepsilon & \lambda\gamma'\Phi \\ \Phi\gamma\lambda' & \Phi \end{pmatrix}
$$

Appendix B: ***Computer package for estimating SEM: LISREL***

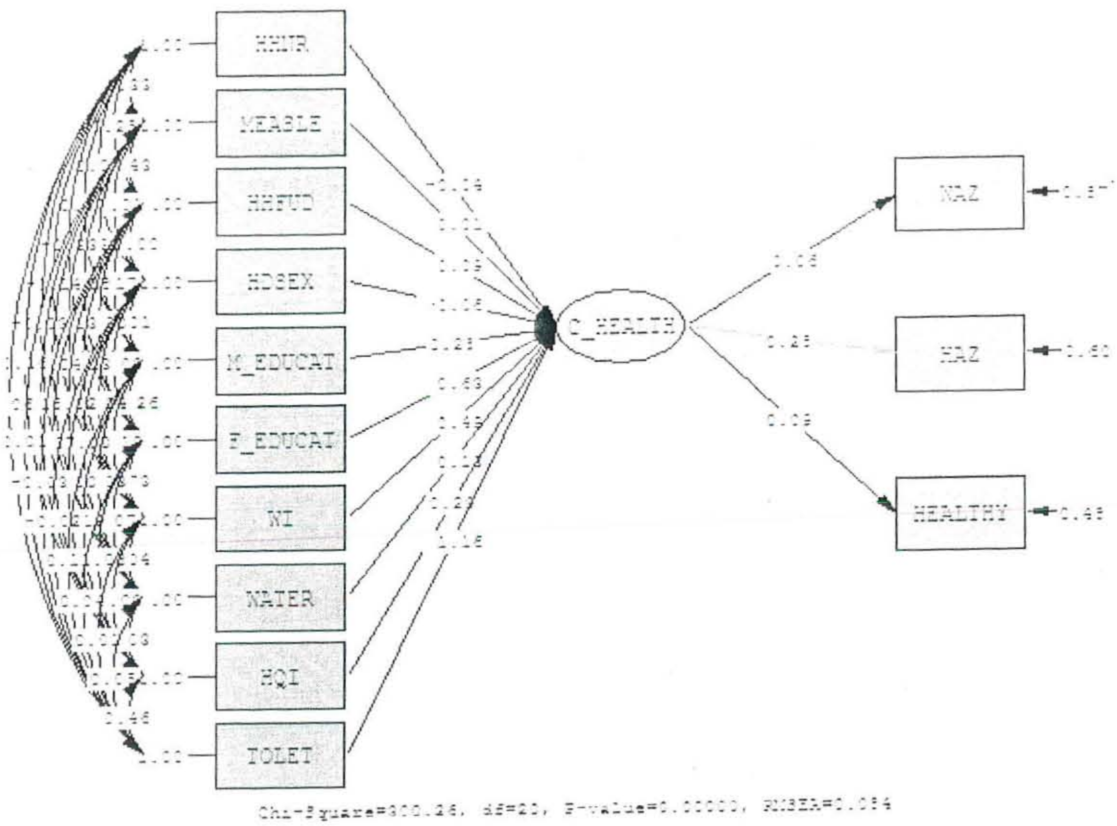
Structural Equation Modeling (SEM) is a relatively recently developed statistical technique based upon factor analysis and multiple regression. When using what might be termed "traditional" statistics data are explored, and relationships tested, to derive a model. When using SEM a model is hypothesized, and the ability of the model to have produced the data is tested. Most commonly the result of that test is a χ^2 value, which gives a probability value for a test of equality between the data and the model.

Recently SEM has seen a surge in popularity both internationally, and in the UK (Tremblay and Gardner, 1996). This has arisen for two reasons, firstly due to recognition of the capabilities and advantages of using an SEM approach, secondly, and possibly more importantly is the availability of user friendly software. LISREL (Linear Structural Relationships), written by Jöreskog and Sörbom, has been around for longer than most other SEM programs, and version 8.80 was released in January, 2009. LISREL is backed up with extensive documentation on Scientific Software's web pages.

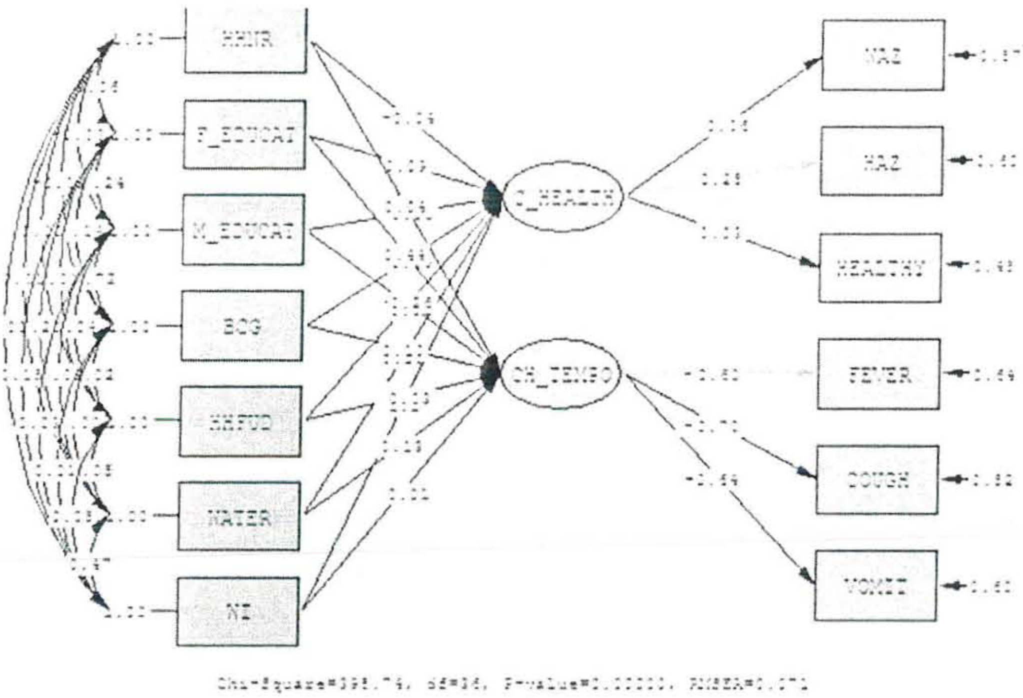
The latest version of LISREL (LISREL-Interactive) is a considerable change for anyone familiar with older versions of LISREL. LISREL comes with two programs, LISREL and PRELIS (PRE-Processor for LISREL). Prelis is used for screening data, and creating the appropriate correlation, covariance and weight matrices.

Both LISREL and PRELIS have undergone considerable changes since their previous releases. In early versions, there was only one way to enter a LISREL model - using LISREL syntax. Even the greatest enthusiast of LISREL syntax would admit that it has a reasonably steep learning curve. The model is estimated by manipulating the elements of 8 matrices, each referred to with a Greek letter (Λ_x, Λ_y, Θ_δ, Θ_ε, Φ, Ψ, Γ and B), to estimate means two additional matrices are required (τ and κ). The LISREL language has two major advantages. Firstly, it is very parsimonious - a model can be written, and communicated, using very little syntax. It is also the most common language for books and technical articles to use.

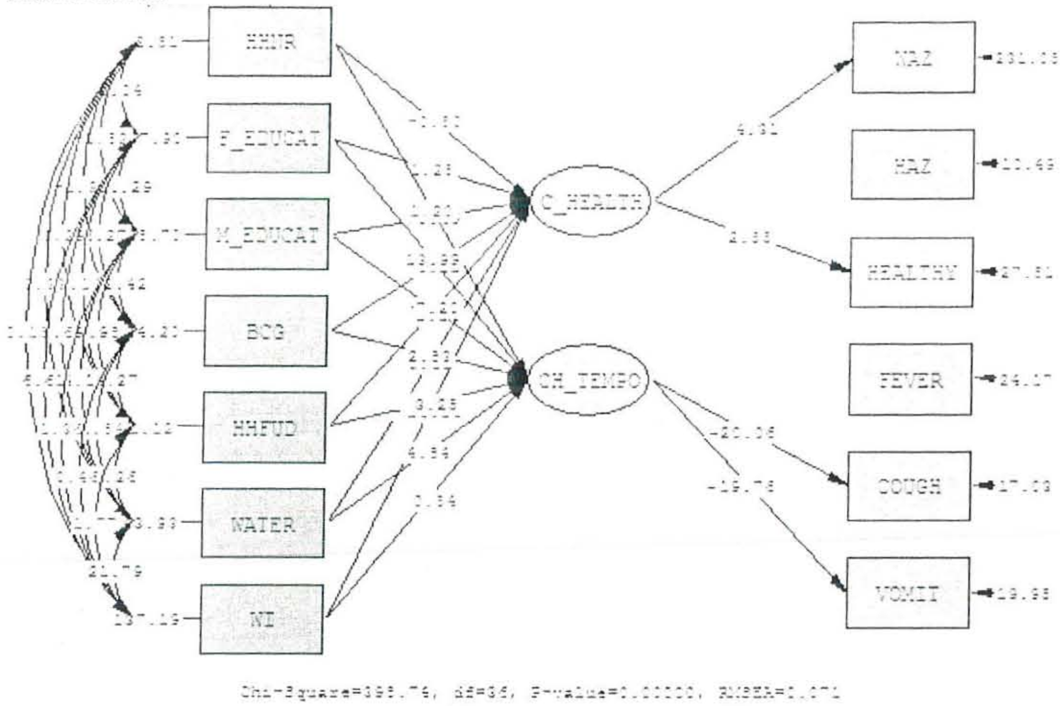
Appendix C: Path Diagram of MIMIC Output for one latent variable



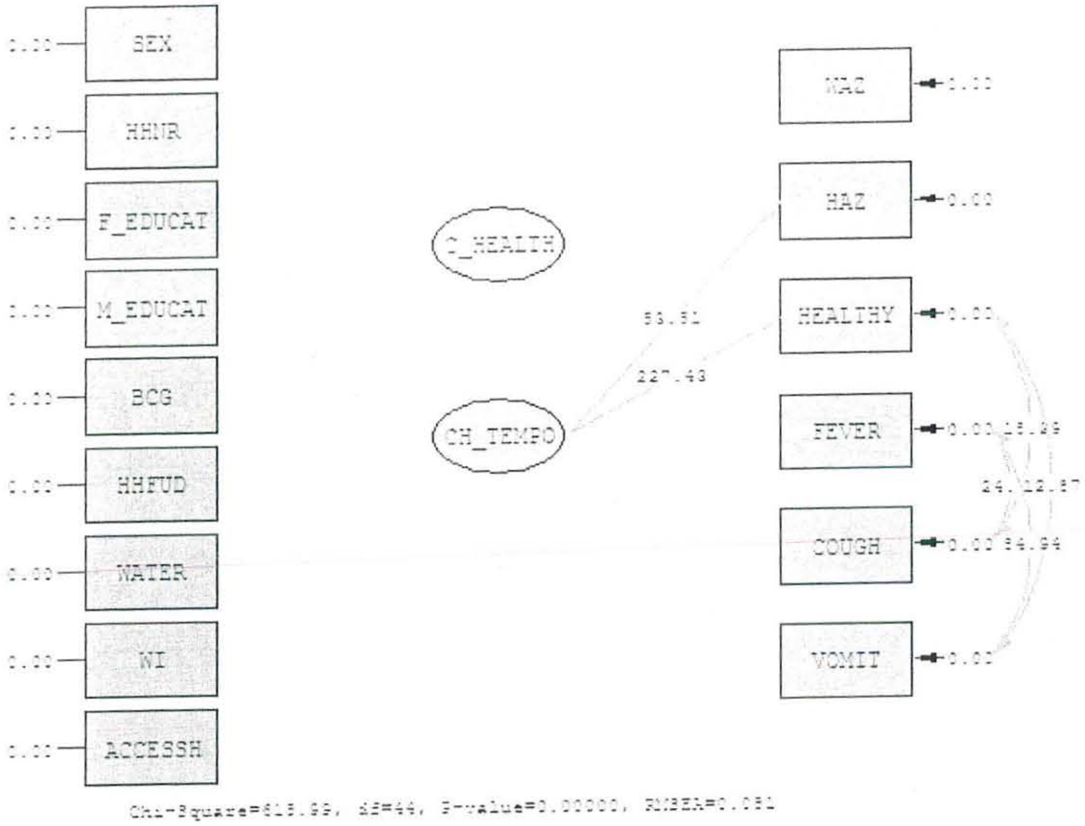
Appendix D.1: Path Diagram of MIMIC Output for two latent variables



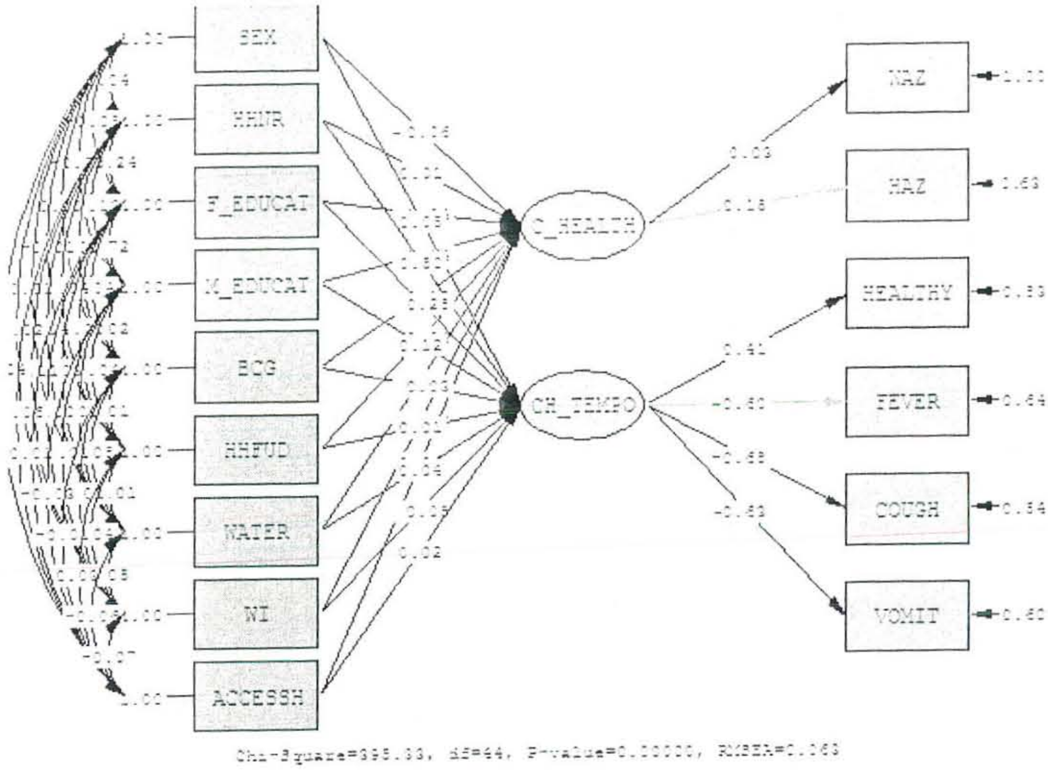
Appendix D.2: The t-values for the two latent variables MIMIC estimate before modification



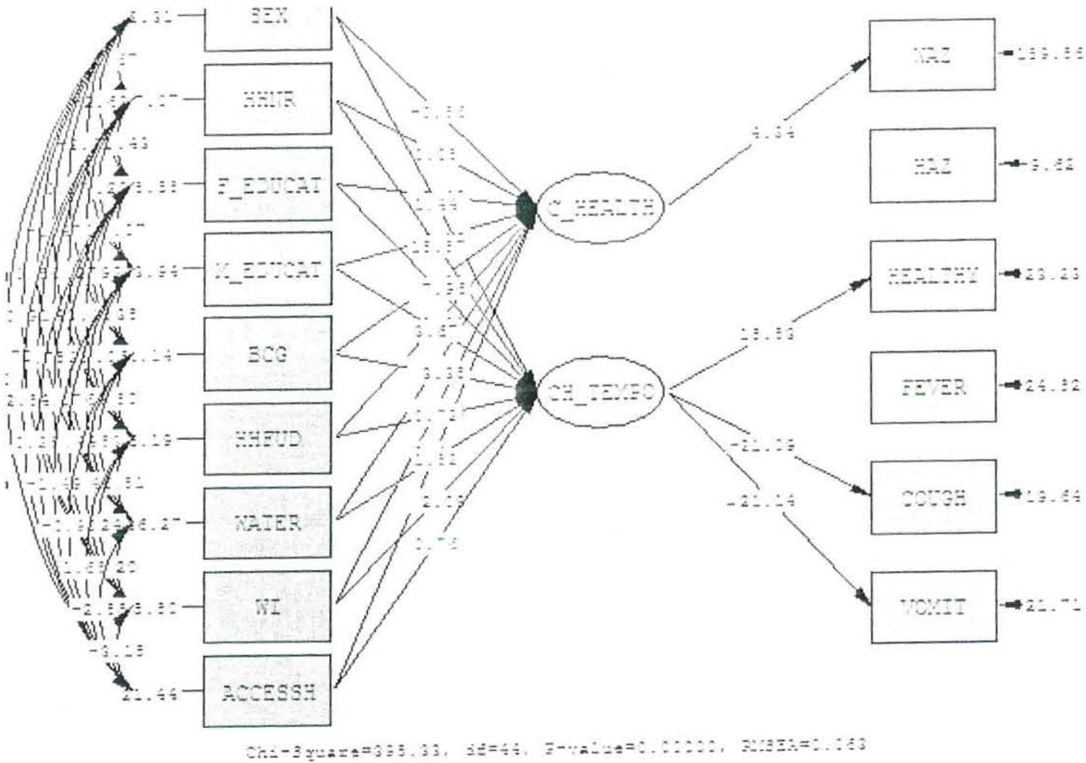
Appendix D.3: Suggested modification for the indicators



Appendix D.4: MIMIC estimates after the modification



Appendix D.5: T-value after modification



Declaration

I, the undersigned, declare that this thesis is my original work and has not been presented for a degree in any other university, and that all source of materials used for the thesis have been duly acknowledged.

The examiners' comments have been duly incorporated.

Declared By:

Name: Biniyam Mekasha

Signature: ______________________

Date: 01/07/08

Confirmed by Advisor:

Name: Tassew Woldehanna

Signature: ______________________

Date: 01/07/2009

Place and Date of Submission: ______________________