USE OF INDICES IN MODELING CHILD NUTRITIONAL STATUS IN MALI

By

Sarah Kaye Light

PLAN B PAPER
Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Agricultural Economics
2005

ABSTRACT

USE OF INDICES IN MODELING CHILD NUTRITIONAL STATUS IN MALI

By

Sarah Kaye Light

While Mali has experienced agricultural growth in recent years, child malnutrition has remained alarmingly high. This thesis examines the contributions that including indices can make in modeling child nutritional status. All indices are derived from three index categories: childcare, sanitation, and feeding. Four indices are constructed through two methods: factor analysis and ad hoc. The models with indices are compared to a preexisting model in order to identify immediate and underlying determinants of height-for-age (HAZ) Z-scores (an indicator of long-term health and nutritional status) in rural Mali and examine the relative impact that changes in these factors will have on average HAZ scores for rural children. Though the original model has a larger number of observations and will, ceteris paribus, have a larger R-squared value, the higher R^2 does not necessarily signify that the model is an accurate reflection of the underlying relationships.

Explanatory power of the individual indices varies. Coefficients of four of the indices are not significant in the regression. Four of the indices are significant and deal with care and feeding. None of the sanitation indices are significant.

ACKNOWLEDGEMENTS

There are numerous people and organizations responsible for contributing to and improving this research. First, I would like to thank the members of my committee, John Metzler, Valerie Kelly, and particularly my thesis advisor John Staatz. Without their guidance and assistance, this research would not have been possible. Jim Tefft has also been a source of tremendous help and information. I also would like to thank USAID, through the Food Security III Cooperative Agreement, and to Michigan State University for their generous support of my studies and research. All remaining errors are the sole responsibility of the author.

To my family and Dr. Frank Dooley: your encouragement and support made this possible. And to Dr. William Masters who first introduced me to African studies. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER 1: INTRODUCTION

1.1 Problem Statement

Childhood malnutrition rates in Mali are alarmingly high. Thirty-five percent of Malian children ages zero to thirty-five months are chronically malnourished, twenty-three percent suffer from acute malnutrition, and forty percent are underweight (Tefft et al., 1999). For children who are mildly malnourished, the risk of dying from a disease is two times higher than that of well-nourished children, moderately malnourished are five times as likely to die, and severely malnourished children have a risk that is eight times higher (UNICEF 1996).

Researchers have a basic idea of what factors contribute directly to malnutrition: insufficient food consumption, disease, and nutrient absorption are among the most commonly mentioned factors (Paknawin et al., WHO). In general, it is believed that increases in agricultural productivity that increase rural incomes will contribute to decreases in malnutrition because parents with higher incomes would be better able to provide adequate food and health care for their children. However, a paradox exists within Mali. Despite steady growth in the Malian agricultural sector during the 1990's, malnutrition did not decline (Penders, 1999). The Malian government now has two divisions dedicated to the implementation of nutrition programs and their coordination across zones. With Mali's political decentralization, local communities now have fiscal authority over their community, leading to greater control over "making decisions that directly affect their lives" (Tefft et al. 2003). Malian policy makers need better guidance on the types of policies and investments that are likely to have the greatest impact on fighting malnutrition, particularly on the relative contribution of immediate determinants (e.g., child's characteristics, child's dietary intake, child's health status), underlying determinants (e.g., household food security, care of mother and child, availability of health services and environment), and basic determinants (e.g., general economic, social, and political environment).

1.2 Objectives

The broad objective of this study is to identify immediate and underlying determinants of height-for-age (HAZ) Z-scores (an indicator of long-term health and nutritional status) in rural Mali and examine the relative impact that changes in these factors will have on average HAZ scores for rural children[1]. The paper will build on work already done on this topic: Castle, Yoder, and Konate's 2001 work on complementary feeding, Christopher Penders's 1999 findings on links between agricultural development and the nutritional status of children, and of course, the endeavors of the Demographic and Health Surveys (DHS). The specific contribution of the paper will be in its methodological approach, which will focus on developing a small number of indices capable of representing complex sets of knowledge, attitudes, and practices about childcare. The usefulness of these indices for modeling the immediate and underlying determinants of nutritional status will be examined.

[1] These standardized scores are calculated as the difference between the individual and the reference population mean, divided by the standard deviation of the reference population.

1.3 Conceptual Framework

As Paknawin-Mock et al. (2000) suggest, "The socio-ecological framework of child growth begins with seven basic needs" – nurturing, cleaning, clean water, medico-spirit care, food, clothing, shelter. Not until these intrinsic needs are met is a child able to grow at its genetic potential (Paknawin-Mock etal, 2000). The conceptual framework guiding this research (Figure 1) has been adapted from Smith and Haddad's (2000) cross-country study on the determinants of malnutrition in developing countries. Following the Smith and Haddad framework, the determinants are classified into three levels: immediate, underlying, and basic. Immediate determinants are the variables that have the most direct effect on the child's nutritional status, underlying determinants are a step further back in the causality (acting on nutritional status through the immediate determinants), and basic determinants are general characteristics of the child's environment that shape both the underlying determinants and the immediate determinants. This study primarily focuses on the immediate and underlying determinants.

The three levels of determinants are interrelated. The immediate determinants are those variables that are directly related to the child. Food consumption (nutrients and calories), and health status (itself determined in part by nutritional status) are represented at this level. These variables are influenced by the underlying determinants, which fall into three broad categories: household food security, mother and childcare, and health environment/services. A household's food security is derivative of its food production and assets. The care received by the mother and child is related to the resources for care.

In this study, two areas of human capital represent the resources for care. The first aspect of human capital considered is quality, which takes into account variables concerning the education of the mother and her personal income. The other aspect of human capital is its availability. For example, the size of the household and its cropping systems will influence the time a mother has to care for her child. The last part of underlying determinants is the health environment and services. The resources available for health include water supply, sanitation, and healthcare services. Basic determinants of a child's health are rooted in the economic, political, and social structure of society. Which zone the child lives in (irrigated rice, cotton, and millet/sorghum) will have an influence on the nutritional status of the child because the economic opportunities and socio-cultural practices regarding women's work and childcare differ across these zones.

1.4 Mali's Agriculture

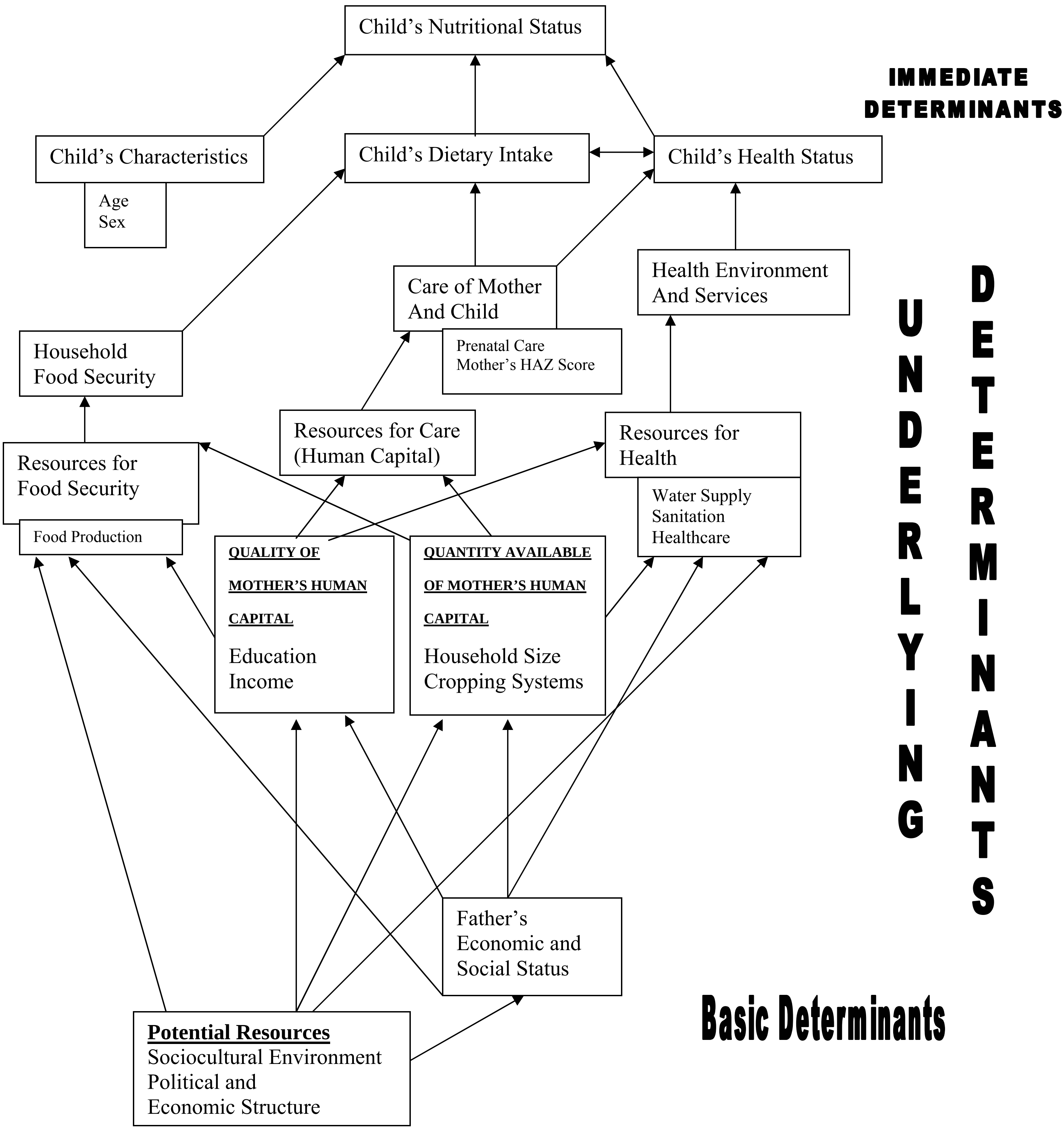
The last fifteen years have been a time of growth for the agricultural sector in Mali. The annual growth rate between 1984 and 1999 for the agricultural sector has been 3.9% (Tefft et al. 1999). The CFA franc experienced a 50% devaluation in 1994, which led to an even more rapid growth rate in the cotton and rice sectors – greater than 9% annually between 1994 and 1999 (Tefft et al. 1999). However, despite the increases in agriculture growth, the DHS II survey indicated that child malnutrition had stayed high since the DHS I survey (Tefft et al. 2001). However, it must be noted that the sampling was different for the DHS I and DHS II surveys, so it cannot be said that child malnutrition worsened between the DHS I and DHS II surveys (Tefft et al. 2000). The DHS II findings showed that 30% of the children surveyed in 1995-96 were stunted and 23%

were wasted (Penders et al. 2000). These statistics on childhood nutrition were found not to be significantly different between regions or between urban and rural zones (Tefft et al. 1999).

1.5 Anticipated Contribution

The implementation of this research project will aid in formulating more precise policy recommendations to combat child malnutrition in Mali. Through the creation of indices, we hope to incorporate a variety of factors into the HAZ model that would otherwise be extremely difficult to capture. If successful, the modeling results will contribute to improved guidelines for alleviating the widespread problem of malnutrition in rural Mali. However, it must be noted that this study focuses solely on the development of the indices. Neither the conceptualization nor implementation of policies or policy guidelines is broached.

Figure 1: Conceptual Framework



CHAPTER 2: CONSTRUCTION AND USE OF INDICES IN REGRESSION MODELS

2.1 Introduction Of The Model; Analytical Methods

This study uses a multiple regression model to explore the determinants of child long-term nutritional status, as indicated by HAZ scores. The model proposed is:

$$HAZ=f(I, U, B, \varepsilon_i) \quad (1)$$

where the immediate determinants (I), underlying determinants (U), basic determinants (B), and random disturbances within the system (ε_i) are used as determinants of a child's HAZ score. This model builds upon the earlier work by Kelly et al. (2004), who analyzed the factors influencing childhood malnutrition in Mali. The analysis here uses the same data set that Kelly et al. used, but rather than just using individual explanatory variables, it constructs indices of several variables, which are included in the model as independent variables. The aim is to examine whether inclusion of these indices improves the ability to "explain" nutritional outcomes, as measured by HAZ.

2.2 HAZ Scores as a Measure of Child Nutritional Status

Many tools are typically used in studies to determine the nutritional status of children. Two anthropometric indicators commonly used for identifying malnourishment are height-for-age Z-scores (HAZ) and weight-for-height Z-scores (WHZ). Both are calculated as the difference between the individual and the reference population mean, divided by the standard deviation of the reference population.

Reference population numbers were taken from the National Centers for Health Statistics/World Health Organization (NCHS/WHO) reference. The measurements are from a "statistically valid random population of infants and children" (COGILL 2003). Though this reference population consists solely of U.S. children, evidence exists that this is an acceptable reference. Studies have shown that the growth patterns of children from varied ethnic backgrounds will be similar for "well-fed healthy preschool children" (Bhandari et al., 2002). From this NCHS/WHO reference, stunting (low height-for-age) and wasting (low weight-for-height) can be quantified and set to a normalized standard represented by the number of standard deviations a child's score is above or below the mean score for the reference population. For example, a HAZ score of zero means the child's height for age is equivalent to that of the reference population mean. A score of –2 means that the child's height for age is two standard deviations below the reference mean. Children are generally considered malnourished when their Z-scores are lower than –2 and severely malnourished if their Z-scores are below –3.

HAZ is an indicator of long-term (chronic) malnutrition and poor health, which results in stunting. WHZ is representative of wasting – a short-term indicator of malnourishment. Low HAZ scores have been correlated with "poor developmental attainment in young children, and poor school achievement or intelligence levels in older children" (Onis et al. 1993). Because a mother's low HAZ score can result in increased obstetric risks, the negative effects of chronic malnutrition can even reach into the next generation (Onis et al. 1993). Identifying the most important determinants of HAZ scores will better aid researchers in making sound policy recommendations for reducing childhood malnutrition.

2.3 The Use of Indices in Regression Models

The aim of adding indices to a model of nutritional outcomes is to gain greater explanation of health influences--i.e., a model that has greater explanatory power. Kelly et al.'s model (Kelly et al. 2004) does not include variables dealing with the aspect of sanitation. This original model attempts to better understand "factors that must be taken into account in the design of programs and policies to reduce malnutrition in Mali" (Kelly et al. 2004). The modified models developed in this study uses Kelly et al.'s model and incorporates eight different variables by way of three categories of indices (childcare, sanitation, and feeding). The indices created are each employed as an explanatory variable. Thus, the indices bring many specific variables into a single comprehensive variable. This then allows the researcher to include more variables in a regression[2]equation than are currently used. Indices can be created to encompass a wider variety of knowledge, attitudes, and practices. This will, theoretically, result in a better estimation of children's HAZ determinants.

The nutrition Knowledge, Attitudes, and Practices (KAP) survey in Mali that generated the data used in this study (see Chapter 3) has a large number of variables. Putting a large number of independent variables in the model raises the risk of multicollinearity, which prevents the researcher from seeing the impact of the individual variables. The

[2] OLS regressions allow the researcher to identify significant variables and thereby assess the impact that small changes will have on the dependent variable

purpose of data reduction, creating an index, is to reduce the number of variables in order to better encompass a full range of practices in one model.

2.4 What is an Index?

An index is a weighted average of variables. The critical issues involved in creating an index include how to pick the variables and how to weight them. The goal of creating an index "is to collapse [a] set of variables linearly in terms of a smaller number of factors" (Faruqee 1979). This can be accomplished in two different ways: a statistical procedure or at the researcher's discretion. The summary of a large number of variables is then put into a regression model.

Organizations such as the International Food Policy Research Institute (IFPRI) and the World Bank have used indices in their research. Arimond and Ruel (2002) created an index for IFPRI using 'ad hoc' methods for analysis or data from the Ethiopia Demographic and Health Survey 2000. A feeding index was derived by summing component scores from variables measuring the following: breastfeeding, bottle use, twenty-four hour dietary diversity, seven-day frequency of feeding, and the twenty-four hour solid/semisolid feeding frequency. Each of these topics was addressed in the survey as a separate question. Points were assigned for each response, depending on the amount of perceived influence on the child's HAZ score. The index varied between zero and nine, with each component given roughly the same weight.

Faruqee (1979) prepared a paper on family planning for fifty countries for the World Bank using factor analysis. The goal of this study was to separate the influences on family planning decisions, using fertility decline as the dependent variable, by analyzing the interrelated variables that represent the different aspects of development. This study took fifteen independent variables and collapsed them into four factor scores. This study allowed the World Bank to reconfirm Liebenstien's economic theory of fertility decline[3] as well as discover underlying factors affecting economic growth (Faruqee 1979).

2.5 Ways of Constructing an Index

There are a variety of different methods for constructing an index. One way is through statistical techniques. Examples include such methods as cluster analysis and factor analysis. These methods allow the data to arrange themselves into factors, but the researcher must decide what each factor represents, giving it a title. Another method of index construction is an ad hoc method. When using this method, the researcher must go through and individually weight each variable.

This study utilizes two ways of constructing an index. Factor analysis is employed in the creation of three groups of indices. Another index is created through an 'ad hoc' method to compare its results with a similar index created through factor analysis.

Factor analysis is a method of explaining a large number of variables with a much smaller number of underlying but unobserved variables. Factor analysis analyzes the arrangements "of relationship among many dependent variables, with the goal of discovering something about the nature of the independent variables that affect them, even though those independent variables were not measured directly" (Darlington et al.

[3] Economic development implies and is accompanied by conditions that induce fertility decline.

1973). When the factors are made and the researcher is viewing the output, it is then that the researcher must decide what the factors represent. The loading factors guide the researcher in his/her conclusions. Nevertheless, it is at this point that human bias enters the study. The 'ad hoc' method relies on human ingenuity rather than statistical techniques. The researcher, for the different components of the index, decides upon scores. This study uses both factor analysis and the 'ad hoc' method.

The Arimond and Ruel (2002) study, which looked at feeding children between six and thirty-six months, provides an example of the construction of an ad hoc index to rate child feeding practices. An example of how points are awarded is the breastfeeding variable. No points were assigned for potential harmful practices (e.g., not breastfeeding). Two points were awarded for breastfeeding between six and twelve months and one point was awarded for breastfeeding between thirteen and thirty-six months. The logic for assigning two points for younger children is that breast milk is more essential for the younger children than the older children. When all of the variables (breastfeeding, bottle use, twenty-four hour dietary diversity, seven-day frequency of feeding, and the twenty-four hour solid/semisolid feeding frequency) are assigned points, they are then summed according to the child's age.

Many difficulties exist when constructing an index through the 'ad hoc' method. First, the index will not work if the point system varies across age groups. Ideally, the same minimum/maximum point range should be used for all age groups. However, this is not always possible when using theoretical/practical logic. For example, children less than six months will sometimes have a greater maximum number of points than the older children. For example, breastfeeding is more imperative for younger children and, therefore, awarded more points when provided to younger children than older children. These points should be balanced with another variable, possibly dietary diversity. Another obstacle with this method is the weighting of the variables. There is no standard method in determining which aspects, for example, of feeding, are more important than others. The weighting scheme in the final equation that is used in this study and in the Arimond and Ruel (2002) study is equal weight for all variables. However, as this is also considered weighting the variables, the author recognizes that it is as arbitrary as giving different weights to the variables.

CHAPTER 3: CONSTRUCTION OF THE INDICES

3.1 Layout of Chapter 3

The intention of this study is to see if creating and including in regression models indices that summarize key independent variables can yield better explanations of nutritional outcomes than only including some of the individual independent variables. Judgment is made by the comparison of a model with indices, as formulated in this study, with a model developed by Kelly et al. (2004).

The following section of this chapter deals with the data and methods used in this study. The study develops indices in three areas for inclusion in the model: childcare practices, sanitation practices, and feeding practices. In developing the indices, variables were chosen for inclusion so that the number of observations is as close as possible to the model run by Kelly et al. (2004). This allows for a valid comparison between the models. Thus, discussion of N (number of observations, or cases) relates more to getting a similar sample to that of Kelly et al. than it does relating to degrees of freedom in the index model.

The method through which the indices were created consisted of three stages. The first step in creating the indices was to select variables for inclusion in an index. This step also provided the researcher a better understanding of the data set. The number of valid cases for a given variable had to be at a minimum of 500. Again, this relates back to needing a good model for comparison with Kelly et al.'s (2004) model. When a variable with a sufficient number of observations was found, a bivariate correlation with the child's HAZ score was run, if the variable was continuous. The continuous variable was then transformed into a binary variable through examination of HAZ means for different responses as well as taking theory into account. Practices perceived as positive actions were given a score of 1 while other practices were given a score of 0. There is one exception to this, in the sanitation index. This variable deals with the mother's hand washing practices, where the top three practices are taken into account and there is a maximum of three points, if the mother uses all practices. If the variable was already a binary variable, a means comparison was run, using 1 Way ANOVA, with the child's HAZ score to determine how the variable responses should be ranked. When all variables with sufficient N were recoded, they were ready for factor analysis.

After the factor analysis had been run, the newly acquired factors were ready to be put into the model as independent variables. The regression was run and the results were compared to Kelly et al.'s (2004) results.

For the feeding index, this study also created an ad hoc index, with the author determining the weighting of the variables. This index was put into the OLS model. The results were then compared to both the factor analysis approach and to Kelly et al.'s (2004) model. The goal of creating an index through factor analysis as well as ad hoc was to determine which approach was more effective.

The last stage of this study is interpreting the results. Conclusions are drawn from the model comparison. Comparison consists of looking at significant variables, comparing the R-squared value and the F statistic. From these results, policy recommendations can be formulated and conclusions drawn about which approach yielded a better index.

3.2 Data And Methods

The dataset used for this research is from the Linkages between Child Nutrition and Agricultural Growth (LICNAG) study, implemented collaboratively by the USAID/MSU III Cooperative Agreement and the Institut du Sahel in 2001/2002. The LICNAG study contains a base socio-economic survey at the household level and three supplementary surveys:

- a Knowledge, Attitudes, and Practices (KAP) questionnaire to collect information about childcare practices from parents of children in the base survey who were under 2.5 years of age
- interviews with the health center personnel to collect information about available services and health center resources, and
- interviews with local government representatives to collect information about Mali's new program of government decentralization and the impacts that it might have on local services (education, health, and general economic infrastructure).

3.3 Historical and Geographical Background

Mali is a country that relies heavily on its agricultural production as a source of national income and employment. This study examines three different areas that depend on three different cropping systems: irrigated rice, millet/sorghum, and cotton/coarse grains. The irrigated rice system that LICNAG studied is in the Office du Niger zone, the cotton/course grain system is in the CMDT (cotton) zone, and the millet/sorghum system studied is in the area of Mopti. To understand the agricultural growth patterns in these different systems, a brief discussion of the history and geography of the different zones.

3.3.1 A Brief Historical Overview

French colonial policy aimed to make Mali major exporter of cotton and the bread-basket of French West Africa. Consequently, During the 1920's France began to emphasize infrastructure development, particularly irrigation, initially in an attempt to grow irrigated cotton in the Office du Niger area, and when that failed, to grow irrigated rice. The first colonization center for irrigated agriculture, built in 1925 along the Niger river, was located in Baguineda (Labaste, 1996). The focus was on cotton and paddy production, but the center was abandoned after a few years (Labaste, 1996). During the 1930's, when the French developed more extensive irrigation systems north of Segou, the Office du Niger (ON) was created to supervise building and operation of the irrigation system, provide extension advice to farm families that were moved into the area to grow crops, and to handle all marketing and processing of the rice produced. Today, the role of the ON is primarily to provide extension and irrigation services, with the private sector handling all the marketing and processing of rice (Tyner et al. 2002).

After the failure of irrigated cotton in the ON zone, French authorities turned to promoting rain fed cotton production in the south-east part of the country, under the leadership of the Compagnie Française de Développement des Textiles (CFDT), which operated throughout West Africa. The company provided farmers inputs on credit and held a legal monopsony on all cotton marketed in French West Africa.

In areas outside of the cotton and rice zones, there was very little colonial investment in agricultural development. In the Mopti area, for example, millet and sorghum production continued using traditional varieties and technologies, largely for home consumption.

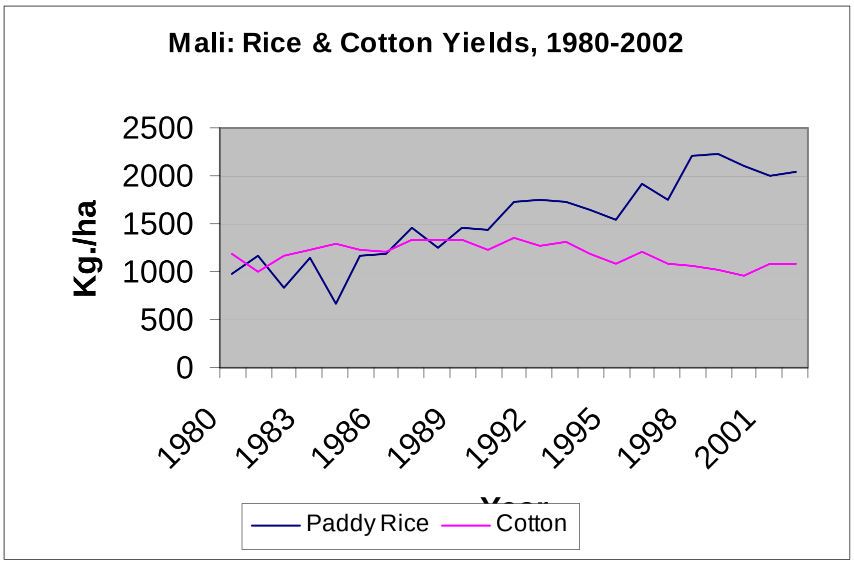
Mali gained independence in 1960, taking "a socialist-oriented approach to economic development" (Tyner et al. 2002). Modeled after the 1928-1959 colonial policies in French Sudan, Mali aimed to expand "the production of export crops" (Bingen et al. 2000). For the first twenty years of independence, the food policies and agricultural development remained relatively unchanged (Bingen et al. 2000). The Malian state took majority ownership position in the cotton company (changing its name to the Compagnie Malienne de Développement des Textiles, or CMDT), but kept the CFDT as a partner and continued similar production and marketing policies for cotton. The state established a grain marketing board, the Office des Produits Agricole du Mali (OPAM), which was granted monopsony marketing rights for all cereals in the country. OPAM was never able to effectively implement its monopsony (especially for cereals other than rice), so a large, technically illegal, parallel trade by private merchants existed alongside the private system, but operated with high transaction costs since it was often harassed by state agents (Dembélé and Staatz, 2002).

The Malian Company for the Development of Textiles (CMDT) put forth strong extension and literacy efforts beginning in the 1960's. These extension efforts focused primarily on agricultural production and very little on health and nutrition. The decade of the 1970's saw a shift to where funding was being directed. In previous years, commodity-based projects got much of the funding, whereas the 1970's focused on integrated rural development programs (Bingen et al. 2000). During the following years, Mali faced a stagnation of food production. This combined with the Sahelian drought led to Mali's food crisis (Bingen et al. 2000). In part in response to agricultural stagnation, the Malian government, with support and pressure from donors, began in in the early 1980s to liberalize agricultural markets in the hopes of creating greater incentives for agricultural production. The reforms began with the Cereal Market Restructuring Program (PRMC). in hopes of increasing farmer's income and cereal production (Tyner et al. 2002, Dembélé and Staatz).

3.3.2 Geography of the Survey Zones

This study focuses on three areas, each with different cropping systems: irrigated rice, millet/sorghum, and cotton/coarse grains. All of these zones have recently experienced an increase in production (Tefft and Kelly 2002). Tefft et al. found that "agricultural growth was highly variable" between production zones (See Figure 3.3.1) and that the demographic group with higher rural income is made up of the most productive farmers.

These authors also documented that the rate of child malnutrition also varies across production systems. The course grains and cotton zones have the highest rates of child malnutrition (as measured by HAZ scores), roughly 45%, while the irrigated rice zone has a lower rate of child malnutrition, about 25% (Tefft et al.).



Taken from Tefft et al.
Figure 3.3.1 Rice and Cotton Yields (1980-2002)

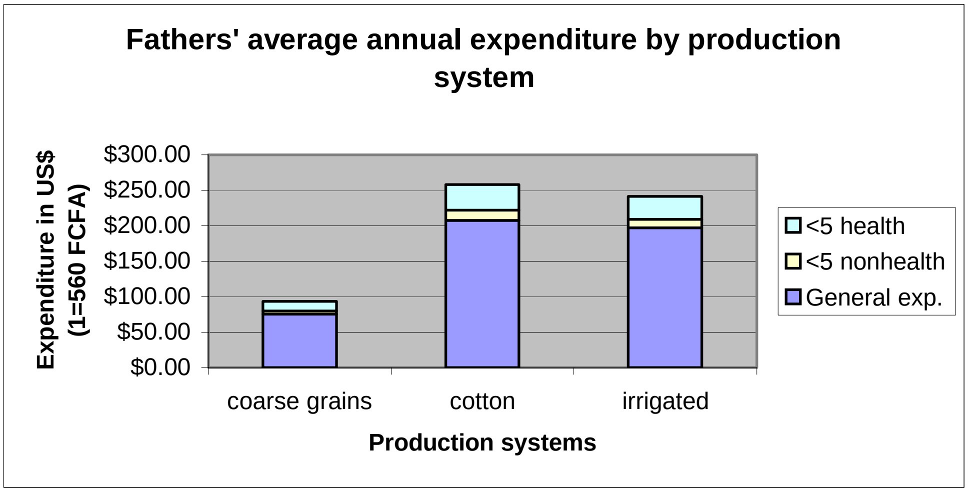
The ON encompasses the Ségou region – Niono and Macina cercles. This is the rice zone, characterized by low climatic risk and moderate price risk (Tefft and Kelly 2002). During the 1990's the aggregate rice production increased from 100 to 300 thousand tons, while yields grew 8% annually (Tefft and Kelly 2002). During the dry seasons, farmers in this zone also produce onions and tomatoes to supplement income (Tefft and Kelly 2002). The manure that comes from livestock is also important to overall income in that it is applied to fields and helps increase yields (Tefft and Kelly 2002). These traction animals serve two three purposes not only in the Ségou region, but also in the other two. The first purpose is to provide the motive power to cultivate fields, thus allowing farmers to expand the area cropped. Second, they serve as an emergency monetary asset.

Thethird purpose, as aforementioned, is a source of organic matter, thereby increasing soil fertility (Tefft and Kelly 2002). The Niono cercle has a large market and good quality transportation infrastructure (Tefft and Kelly 2002). Therefore, since the 1980's Niono has become a commercial center for traders coming from Bomako (Tefft and Kelly 2002). Mancina, on the other hand, has been historically difficult to access. However, since the 1994 devaluation, traders have started to seek out this village rather than making these farmers transport their production to more central locations (Tefft and Kelly 2002).

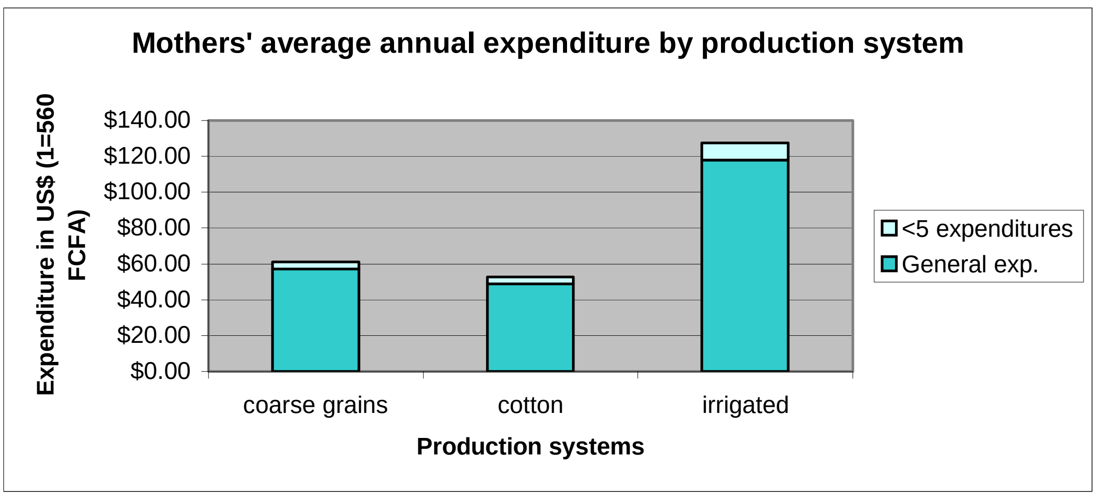
The Sikasso region is in the cotton production zone. The cercles that were chosen to represent this area are Koutiala and Kolondieba. This zone has more climate risk and less price risk than what is found in Ségou (due to the lack of irrigation); this is a zone of intermediate agricultural growth (Tefft and Kelly 2002). Between 1992 and 1997, annual production rose 18% while yields declined 2% (Tefft and Kelly 2002). As in the Ségou region, traction animals are important, as are other crops (Tefft and Kelly 2002). The Sikasso region produces not just cotton, but also relies heavily on sorghum and maize (Tefft and Kelly 2002).

The Mopti region is the coarse grain (millet/sorghum) production zone. The cercles that were chosen to represent this area are Koro and Bandiagara. Mopti is less market oriented than the other regions (Tefft and Kelly 2002). It is characterized by high climate and price risks with no consistent pattern of recent growth neither yields nor production (Tefft and Kelly 2002). During the dry season Mopti produces onions by constructing small dams (Tefft and Kelly 2002). Some of this region's farmers attempt growing other vegetables, but lack of water is often a problem and most wells dry up (Tefft and Kelly 2002). Just as in the other two regions, livestock is an important aspect of rural life. However, the average farmer in Mopti has fewer livestock than the average farmer in either the Ségou or the Sikasso region (Tefft and Kelly 2002). Recent road improvement in the Bandiagara cercle has attracted many small truckers (Tefft and Kelly 2002). This makes marketing onions and purchasing cereals easier for the farmers. However, Koro is still isolated due to the second rate roads leading to the markets (Tefft and Kelly 2002).

In part because of differences in agricultural productivity, average per capita incomes also vary by zone. Since income is not easily measured, in this study total expenditure is used as a proxy for income. Not only does income vary between production zones, but also between mothers and fathers in those production zones. (See Figure 3.3.2 and Figure 3.3.3). These figures separate out expenditure on children under five years. Contrary to common thought, fathers spend more on their younger children than the mothers do. According to the 1995 DHS II survey, "the within-household distribution of additional income from the best performing crop sectors favored increases in discretionary consumption spending by individuals rather than increases in general household expenditures on basic foods and health care" (Tefft and Kelly 2002).



Taken from Tefft et al.
Figure 3.3.2 Father's Income by Production Zone



Taken from Tefft et al.
Figure 3.3.3 Mother's Income by Production Zone

3.4 Survey Implementation

The data for testing most of the hypotheses come from the KAP and household surveys. The household survey focuses on particulars involved in household operation and maintenance (demography, education, standardized height, assets, expenditures, cereal consumption, inputs and outputs for agricultural production activities, non-farm income). The 750-household sample was selected to represent three types of production systems found in Mali (irrigated rice, cotton, and millet/sorghum), with each zone represented by 250 households. The irrigated rice zone used in the study is in the area of Niono, north of Ségou (see Figure 3.4.1); the cotton zone is around Koutiala, north of Sikasso; and the millet/sorghum area studied was near Bandiagara, east of Mopti. Administrative units (two cercles[4] per production system) were selected purposively to represent different levels of infrastructure and access to markets. Villages were selected within each cercle from among the villages in the 2001 DHS sampling frame.[5] Households were selected randomly among all households in the DHS sampling frame having at least one child less than 4 years of age at the time the survey began.

[4] A cercle is a Malian administrative unit roughly equivalent to a county in the U.S.
[5] The LICNAG survey used a subset of the DHS sampling frame in an effort to make the two different types of surveys complementary.

Figure 3.4.1 Map of Mali

Base 801725 (R00787) 12-94

The KAP questionnaire was designed to assess how well the parents of children younger than 2.5 years of age knew and followed childcare practices prescribed by the World Health Organization. Enumerators questioned these mothers (N=811) and fathers (N=735). Interviewers also conducted open-ended discussions with parents about the initial household survey results in hopes of gaining more information about factors affecting childcare (Wise et al., 2002). This was followed by a structured questionnaire to collect the KAP information.

Data available on children's characteristics include age in months, gender, and relationship to the head of the household, plus monthly anthropometric measures and information on morbidity. Included in the category of household characteristics are: mother's and father's age, standardized height, and education; number of people living in the household; information about number and quality of household assets (livestock, land, buildings, vehicles); and expenditures (collected for all adults using fortnightly interviews throughout the year). Community characteristics include distance to health facilities, number of grades taught in local schools, and the agricultural zone (cotton, rice, or millet/sorghum).

Though a considerable range of data is available, there is also a large amount of missing data. The KAP sample represents a small subset of the larger sample covered by the household survey. The variables chosen for inclusion in the indices should have a large enough number of observations so that the total number of cases used in estimating the model with indices is roughly the same as that in Kelly et al.'s (2004) model.

Before selecting the variables to use in the factor analysis, tests were run to determine which variables were suitable for inclusion in an index. The first step to determine which variables are viable for factor analysis is to look at the number of observations. When a variable with an acceptable number of observations is established, the next step is to run significance tests of the bivariate correlations (or 1 way ANOVAs) between the explanatory variables and the HAZ scores. T-tests are run when a question has two responses. If there are multiple responses to a question, the answers are again compared with the HAZ in order to find the means of the answers. From there, the question is transferred into a binary variable, or, in one case, a truncated variable for use in factor analysis. This is done through grouping the responses with either similar HAZ scores, or theoretically similar practices.

If a variable has a sufficient number of observations and is correlated with mean HAZ scores, then it is selected as a candidate for inclusion in an index. The indices were developed through factor analysis or, in one case, on an ad hoc basis, as described below. Then they are put into a model and then regression analysis is run to test their significance.

This study uses factor analysis to create indices in three areas: childcare practices, sanitation, and feeding practices. Including an index through factor analysis in the regression analysis helps avoid the problem of multicollinearity[6]. This is a good process

---

[6] Multicollinearity is a term that refers to high correlation among the independent variables in a multiple regression model that prevents the researcher from seeing the independent effects of any one of the

for index creation when there is no a priori knowledge of what weights should be assigned to the variables going into the index. The first step in creating a factor analysis index is to choose a broad set of variables that are potential candidates to include in the index, after which the factor analysis picks out the most important variables that go into a factor. Factor analysis allows the variables to weight themselves in the index (including potentially a weight of zero). These weights, called factor loadings, show how the variables all relate to an unobserved common factor.

In addition to the factor analysis, one index, a feeding index, was also created through individually weighting the variables. The goal of creating an improvised index was to see if it could better explain its variables than those indices created through factor analysis. A similar index was created by IFPRI (Arimond and Ruel 2002) and, thus, used as an example. The KAP survey has many of the same variables that are used in the IFPRI index.

3.5 Child Care Indices

3.5.1 Hypothesis

We hypothesize that mothers who seek medical care both for themselves and their child will have better nourished/healthier children than mothers who do not seek medical care. This hypothesis takes into account the resources for human care that influence the care of the mother and child, thereby affecting the health status of the child and, finally, the nutritional status of the child. An Indonesian study by Paknawin-Mock et al. (2000)

---

variables; it is usually suspected when some correlations are "large," but an actual magnitude is not well-defined (Wooldridge 2000) found that “health-oriented activities appear to have a worthwhile impact” on child growth. Preventative health care has also been shown to increase child growth (Cebu Study Team, 1991). Kebede (2003) showed that low birth weight is a reflection of “the standard of mother-and-child-care practices” as well as a reliable “predictor of wasting prevalence for children under five in Asia, Africa, and Latin America”. Adair (1997 for the Philippines) concluded that the probability for stunting to occur in a child is inversely correlated to their birth weight. However, this data set does not include birth weight of the child, thereby not allowing this study to examine the relationship between birth weight and HAZ. For the proposed study, the following areas will be tested: health center use, prenatal care of the mother, and type of care given when the child is ill.

3.5.2 Variable Descriptives

Table 3.2.1 presents descriptive information on the variables that were considered for inclusion in the factor analysis for childcare practices. The prenatal care variable was posed in the survey as a yes/no question. With a total of 806 valid responses, the 59% of mothers received prenatal care and had significantly (p=0.001) healthier children than those mothers who did not obtain prenatal care.

Information is also available on where the mother gave birth to the child. Mothers who gave birth in a health facility, as compared to elsewhere, had significantly healthier children (N=799 total respondents, p=0.013).

The only continuous variable used as a candidate for the care indices is the number of times the mother had visited the health center in the last twelve months. This variable is recoded into a categorical variable with two categories: no trips to the health center (66%) and one or more than one trip to the health center (34%). There are a total of 699 observations. Mothers who had visited a health center had children with significantly higher HAZ scores than those who did not ($p= 0.002$).

Another variable included as a possible component in one of the care indices is where the mother seeks medical care when she is not pregnant. Those responses concerning a health related facility were grouped with mothers who were never sick and all other responses were grouped into the same category. This variable has N=792 observations, and mothers who sought care at health facilities had children with significantly higher HAZ scores ($p<0.0005$).

One survey question asked the mother what prevented her from assuring a good life for her child. The new variable was grouped into two categories: no constraints and various constraints. Mothers who saw no constraints had children with significantly higher HAZ scores than mothers who cited constraints ($p< 0.00005$).

The variable addressing to whom the child was taken for its last episode of diarrhea combines two categorical variables. Mothers who responded that they took their child to the health center are recoded with those mothers whose children are very young and have never had diarrhea, as the latter response suggests that the mothers are taking effective preventative actions. The other category for this variable encompasses all other responses (traditional healer, grandparent, etc…). This variable yields 688 usable observations and is significantly correlated to mean HAZ scores at the p=0.006 level.

The final variable included as a possible component in the care indices concerns what type of treatment is given to the child during episodes of diarrhea. When the HAZ score means are examined with each category of responses, theory and application are employed in order to recode the responses into two categories for the factor analysis. Treatment appearing to have a positive effect on the child's nutritional status, such as an SRO[7] packet or never having diarrhea, are grouped together. Likewise, actions judged (either on the basis of a priori knowledge or through examination of the data) to have a negative effect on the child's nutritional status are grouped together. This grouping yields a variable that the one-way ANOVA shows to have a significant correlation with mean HAZ scores (p=0.001, N=686).

[7] Solution de réhydratation orale (oral rehydration solution): a packet of salt and sugar to mix with water in an effort of rebalancing electrolytes and providing energy

Figure 3.5.1 Descriptive Statistics of Variables Used as Possible Components of Care Indices

Variable Name	Variable Description	No. of Obs.	Min	Max	Mean	Relationship to HAZ
Prenat	Dummy=1 if mother had 1 or more prenatal visits, 0 otherwise	806	0	1	0.59	Mean HAZ Prenat-=0 -1.93 Prenat=1 -1.57 t-test significant @ 0.001
LocDeliv	Dummy = 1 if mother delivered in a health facility, 0 otherwise	799	0	1	0.29	Mean HAZ LocDeliv-=0 -1.79 LocDeliv=1 -1.51 t-test significant @ 0.013
Mothcare	Dummy = 1 if mother's general health care during previous year was at a health facility or if mother not ill, 0 otherwise	792	0	1	0.58	Mean HAZ Mothcare=0 -2.01 Mothcare=1 -1.49 t-test significant @ <0.00005
Movisit	Dummy = 1 if mother visited health center one or more times in previous 12 months, 0 otherwise	699	0	1	0.34	Mean HAZ Movisit=0 -1.94 Movist =1 -1.56 t-test significant @ .002
Perccons	Dummy = 1 if mother perceives no constraints to providing good child care, 0 otherwise	795	0	1	0.17	Mean HAZ Perccons-=0 -1.79 Perccons=1 -1.29 t-test significant @ <0.0005
Sdiaadv	Dummy = 1 if mother sought advice re diarrhea treatment from a healthcare provider or if child had no diarrhea, 0 otherwise	688	0	1	0.54	Mean HAZ Sdiaadv-=0 -1.86 Sdiaadv=1 -1.54 t-test significant @ 0.001
Treatdia	Dummy = 1 if treatment for diarrhea was modern (SRO or prescription medicine), 0 otherwise	686	0	1	0.67	Mean HAZ Prenat-=0 -1.98 Prenat=1 -1.55 t-test significant @ 0.002

3.5.3 Omitted Variables

The questionnaires included many variables that were potentially very interesting for inclusion in the child-care indices but which had to be excluded because there were not enough valid observations on them to allow the total sample size to be comparable to that of Kelly et al.'s analysis[8]. In this section, we briefly discuss some of these omitted variables.

In this study, the person who takes on the role of primary caregiver for a child was expected to be an important element in one of the childcare indices. However, upon examination of the primary caregiver variable, results show that neither the principal caregiver nor their age was significantly correlated with the child's HAZ score.

Another aspect of childcare expected to affect the child's nutritional status, but which did not, was whether and how frequently the child was left under the supervision of someone other than the mother herself. The age of the child itself when the mother begins leaving it with another person was not significantly related to the child's HAZ score, nor was the number of times per week the child was left with another caregiver.

One more variable not included in the factor analysis was whether the child was left with a young family member. If the mother had to leave the child with a young female family member (fifteen years and under), leaving the child one time per week has a positive impact on the child's HAZ score. This variable giving rise to this surprising result was

[8] The Varimax rotation with Kaiser Normalization was used. The principle components method was used for extraction. The regression method was used to create the scores.

neither used in the factor analysis for the care indices nor in the model itself, however, due to the small number of observations.

The final set of child-care observations concern help that the mother received before, during, and after pregnancy. The child's father was asked these questions. Those fathers who replied 'yes' to having helped their child's mother with her jobs during her last pregnancy have children with lower HAZ scores than those father's who did not. Similarly, fathers who paid a domestic to help their wife during her last pregnancy have children with an average HAZ score much lower (-2.7050) than those whose father responded 'no'.

3.5.4 Factor Analysis

Factor analysis of the interrelated variables used to form the care indices shed light on their cluster association. By examining the coefficients of the variables for each component, we are able to discern what each factor best represents. (See Figures 3.2.2 and 3.2.3)

In the rotated component matrix, each component corresponds to a child-care index. Each index is independent of (orthogonal to) the other indices (components) in the matrix. Component 1 is highly correlated with the following variables: where the child was taken for last episode of diarrhea and treatment for last diarrhea. Thus, from the factor analysis output, we interpret this component to represent the care of and treatment for a child during episodes of diarrhea. Where the child was taken for its last episode of diarrhea has a factor loading score (correlation coefficient with the underlying factor) of 0.940. Treatment for the child's last episode of diarrhea has a factor loading score of 0.936. Component 1 accounts for roughly 33% of the variance in the data.

Component 2 has two variables with high factor loading scores. The numbers of times the child was taken to a health center during last 12 months and where the mother gets medical care (other than prenatal care) account for about 22% of the variance in the data. Component 2 is thus taken to represent the mother's usage of the health facilities. The number of times the health center was visited during last 12 months has a factor loading score of 0.804, and where the mother gets medical care (other than prenatal care) has a factor loading score of 0.768.

Two variables have high factor loadings on component 3. The variables are where the mother gave birth to the child and the prenatal care she received. The factor loading scores are 0.925 and 0.595, respectively. Therefore, we interpret Component 3 as representing prenatal and delivery care.

Figure 3.5.2 Care Index Matrix: Factor Loadings on Principal Components

Rotated Component Matrix (a)

	Component		
	1	2	3
No. of times to health center during last 12 months	.056	.804	.131
Where child was taken for last episode of diarrhea	.940	.056	.062
Main constraints against assuring good life for child	.038	.124	.064
Treatment for last diarrhea	.936	.121	.020
Prenatal care	.013	.478	.595
Where gave birth	.063	.058	.925
Where get medical care (other than pregnancy)	.130	.768	.089

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 5 iterations.

Figure 3.5.3 Care Index Variance: Total Variance Explained

Component	Initial Eigenvalues		
	Total	% Of Variance	Cumulative %
1	2.31	33	33
2	1.51	22	55
3	.94	13	68
4	.79	11	79
5	.69	10	89
6	.55	8	97
7	.21	3	100

3.6 Sanitation Indices

3.6.1 Hypothesis

We hypothesize that mothers who follow proper hygiene and sanitation practices will have better nourished/healthier children than mothers who do not follow these practices. Ruel et al. (1999) found that hygiene practices are strong determinants in the nutritional status of children in the population of Accra. Adams et al. (2002) found that Malian children between the ages 1-5 years are more susceptible to environmental factors such as poor hygiene than are post neo-natal children. In this analysis, information about sources of drinking and cooking water, disposal of animal and infants' feces, number and type of latrines, hand washing practices, and usage of mosquito nets will be used to examine links between hygiene and sanitation practices and child nutritional status.

3.6.2 Variable Descriptives

The first category of sanitation variables examined includes variables reflecting the household's actions to protect themselves against illnesses, such as malaria, that are transmitted by mosquitoes. Mothers were asked if they and their child slept under a mosquito net the night before the survey. Answers to both questions were virtually identical; 38 percent responded yes. In comparing the mean HAZ scores of children having slept under nets to those who did not, we found a statistically significant difference ($p<0.00005$), with average scores for those using nets being -1.89 vs. –1.41 for those not using nets (Figure 3.3.1 at section end). One possible explanation for these results is that people are more likely to use mosquito nets in areas where mosquitoes are more prevalent. These same areas may also be where children are contracting malaria and other insect-borne diseases.

Where the mother disposes of the child's feces also has a significant ($p<0.0005$) association with the child's HAZ score. Feces disposal in latrines (which was practiced by 20% of the sample) had a positive effect on the child's HAZ score compared with fecal disposal elsewhere (behind the house, in the streets, in the river).

Child and mother hand washing practices were also included as potential candidates for inclusion in the sanitation indices. The child's hand washing practices only concern whether or not his or her hands are washed before meals; of which 96% responded that they do. The mother's hand washing habits were approached in a somewhat different manner. Mothers were asked which times of the day they wash their hands (before/after preparing meals, after using the toilet, etc…). The responses with the highest mean HAZ scores contribute to scoring this variable. The mothers are awarded points according to how many of these good practices they employ.

The questionnaire included four variables dealing with the household's water sources: water source for cooking in the dry season and in the rainy season, and water source for drinking in each of the two seasons. These variables were also used in the sanitation factor analysis index. A binary variable for the water source was constructed as follows: covered wells, both public and private, were combined with tube wells to form one category. All other water sources (the river, irrigation ditches, streams) comprise the other group. Groupings were devised in this manner across all four variables because it was shown that children getting their water from covered and tube wells, both public and private, had significantly higher ($p<0.00005$) HAZ scores than children who obtained water elsewhere. (See Figure 3.3.1)

Figure 3.6.1 Sanitation Index Variable Descriptives

Variable Name	Variable Description	No. of Obs.	Min	Max	Mean	Relationship to HAZ
Dry Season Cooking Water	Dummy=1 if private/public covered well/tube well, 0 otherwise	803	0	1	0.41	Mean HAZ DryCook-=0 -1.84 DryCook=1 -1.53 t-test significant @ 0.004
Dry Season Drinking Water	Dummy =1 if private/public covered well/tube well, 0 otherwise	802	0	1	0.44	Mean HAZ RainyDrink-=0 -1.87 RainyDrink=1 -1.49 t-test significant @ 0.001
Rainy Season Drinking Water	Dummy =1 if private/public covered well/tube well, 0 otherwise	792	0	1	0.39	Mean HAZ RainyDrink-=0 -1.84 RainyDrink=1 -1.51 t-test significant @ <0.002
Rainy Season Cooking water	Dummy =1 if private/public covered well/tube well, 0 otherwise	806	0	1	0.37	Mean HAZ RainyCook=0 -1.84 RainyCook =1 -1.49 t-test significant @ 0.001
Feces Disposal	Dummy = 1 if feces disposal in latrine, 0 otherwise	802	0	1	0.20	Mean HAZ FecDisp-=0 -1.83 FecDisp=1 -1.30 t-test significant @ <0.00005
Mosquito Net Use	Dummy = 1 if child used mosquito net last night, 0 otherwise	801	0	1	0.36	Mean HAZ MosqNet-=0 -1.89 MosqNet=1 -1.41 t-test significant @ <0.00005
Mother Handwash Practices	Possible of 3 good hand washing practices	806	0	3	0.30	Mean HAZ MoHand-=0 -1.73 MoHand=1 -1.63 MoHand=2 -1.55 MoHand=3 -1.83 Not Significant
Child Hand washing Before Meals	Dummy=1 if child washes hands before meals, 0 otherwise	793	0	1	0.96	Mean HAZ Prenat-=0 -1.81 Prenat=1 -1.70 Not Significant

3.6.3 Omitted Variables

Sometimes a significant variable is omitted from inclusion in an index. In this study, the mothers were asked if they slept under a mosquito net and if their child slept under a mosquito net. For both variables, the positive responses are significant at $p < 0.00005$. When the variables are compared to each other, it is found that the respondents gave the same answer for each question, 38%. Including both variables in the factor analysis is redundant and is not productive in bettering the index itself, as the matrix will not invert.

Other variables were omitted from inclusion due to lack of sufficient observations. For this reason, none of the questions from the father's survey are used. Fathers responded to questions such as: do they purify water in their household to make it potable, how do they purify the water, and where do they keep their animals during the night. However, none of these questions yield enough observations for inclusion in the index.

3.6.4 Factor Analysis

From the five variables employed in the factor analysis for the sanitation indices, four components with acceptable Eigenvalues are produced. Component 1 accounts for nearly 47% of the variance in the data and Component 2 yields 15% of the variance in the data. Components 3 and 4 represent 13% and 12% of the variance in the data. (See Figure 3.3.3)

Component 1, representing the first sanitation index, has four variables with high factor loading scores. The four water components are the high loading variables. All of these variables have factor loading scores over 0.93. Therefore, Component 1 represents potable water sources. (See Figure 3.3.2)

Component 2 has two variables with high factor loading scores. If the child used a mosquito net the previous night produces a factor loading score of 0.822. Where the mother disposes of the child's feces has a factor loading score of 0.813. Thus, Component 2 represents avoidance of fecal/mosquito borne diseases. (See Figure 3.3.2)

Components 3 and 4 are very similar. Component 3 represents the mothers' hand washing practices. It has a factor loading score of 0.999. Similarly, Component 4 represents the child's hand washing practices. It, too, has a factor loading score of 0.999.

Figure 3.6.2 Sanitation Index Matrix: Factor Loadings on Principal Components

Rotated Component Matrix (a)

	Component			
	1	2	3	4
Disposal of feces from children not using toilet	.094	.813	.012	2.178E-05
Child used mosquito net last night	.060	.822	.039	-.020
Water source for cooking for rainy season	.932	.148	.037	-.044
Rainy season drinking water	.936	.162	.030	.035
Water source for cooking for dry season	.944	.097	.083	.038
Dry season drinking water	.942	.124	.081	.040
Mother's use of hand washing practices	.032	.029	.999	-.008
Practices of washing the child's hands	.011	-.016	-.008	.999

Extraction Method: Principal Component Analysis. Rotation Method: Quartimax with Kaiser Normalization.
a Rotation converged in 3 iterations.

Figure 3.6.3 Sanitation Index Variance: Total Variance Explained

Component	Initial Eigenvalues		
	Total	% Of Variance	Cumulative %
1	3.74	47	47
2	1.25	16	62
3	1.00	13	75
4	.98	12	87
5	.65	8	95
6	.25	3	98
7	.11	1	99
8	.03	0.32	100

3.7 Feeding Index

3.7.1 Hypothesis

WHO and UNICEF have jointly devised a set of feeding guidelines that are based on the age of the child. We hypothesize that mothers who follow feeding practices recommended by the WHO/UNICEF have better nourished/healthier children than mothers who do not follow these practices. These guidelines were formulated with the changing nutritional requirements of growing children in mind so that proper nutrition is, theoretically, attained at every age, thereby allowing the children to reach their genetic growth potential. For example, exclusive breastfeeding is recommended for the first six months of life, after which complementary feeding with a variety of recommended foods is begun in addition to breastfeeding through age two. We will test this hypothesis using data on breast-feeding practices (frequency, duration, etc.) and the diversity of complementary foods given to children of different ages.

Two feeding indices are created for this study: one through factor analysis and the other through an ad hoc method. The purpose of creating the feeding index through two methods is to compare the final regression results. In this manner, we can see which method does a better job in accounting for the impact of feeding practices on HAZ scores.

3.7.2 Variable Descriptives

Factor Analysis Feeding Index

One of the variables considered for inclusion in the factor analysis is whether or not the mother breastfeeds her infant. An interesting feeding detail is that mothers who were not currently breastfeeding, 28% of our sample, tended to have children with higher HAZ scores, -1.32, than mothers who were currently breastfeeding, -1,84. Again, this goes against theory and the WHO's feeding recommendations and yet is significant at the $p<0.00005$ level. From further analysis, this variable is related to feeding frequency. Children being fed solid food at high frequencies also tend to be the ones not being breastfed. One possible explanation of this is that mothers without the time or the ability to breastfeed, possibly due to late or no milk supply prefer to feed their children solid foods. This gives the child a larger variety of nutrients, vitamins, and minerals. In line with this explanation is that mothers who do not have the financial resources or time to feed their children solid foods will breastfeed longer.

Another variable included for factor analysis is whether or not the mother believes her child, if that child is less than six months, is well nourished. Ninety percent of the mothers responded that they do think their infant is well nourished. Their children have an average HAZ score of –1.64, as compared to the –2.30 of those mothers who do not think their child is well nourished. This knowledge variable is significant at $p<0.00005$.

Similarly, if the child is receiving complementary feeding, mothers were also asked if they were content with their child's nutritional status. Seventy-nine percent said that they are content with their child's nutritional status. They have children with a higher average HAZ score, -1.59, than mothers who are not content with their child's nutritional status, -2.13. This attitude is significant at $p<0.00005$, possibly showing that mothers are aware when their child is not well nourished.

Mothers were also asked if they continue to breastfeed after beginning complementary feeding. Mothers who do continue to nurse have children with lower HAZ scores than mothers who stop nursing: -1.75 and –1.35, respectfully. This difference is significant at the 0.005 level. One explanation for these findings could be that it is typically the poorest mothers who continue to breastfeed, giving that child mostly breast milk.

Also included in the factor analysis is the number of food products a child eats each day. Eighty-five percent of the children consume two or more food products on a daily basis. Children who eat fewer than two food products have an average HAZ score or –2.13, compared to the –1.62 average HAZ score of children who eat two or more products. This is significant at the 0.002 level.

Ad Hoc Feeding Index

The first variable included in the feeding index is the age at which solid complementary food is introduced. The survey gives a continuous variable, which this study transforms into a binary variable. One point is awarded for introducing solid complementary foods at either five or six months. Otherwise, zero points are awarded. T-tests are run to compare children who started consuming solid foods at the recommended age, five and six months, with those children who started eating solid foods earlier or later.

Another variable in this index is a diet diversity variable. Milk, animal proteins, lipids, and legumes comprise this variable. One point is awarded for eating animal products in the last twenty-four hours, one point is given for eating fats, half a point is given for consuming legumes, and half a point is also given for drinking milk. Running a correlation between this dietary diversity variable and HAZ shows a significance of $p<0.00005$ and a Pearson's correlation of 0.164.

The third variable in the feeding index represents actions taken when the child refuses food. HAZ means for the individual responses as well as theory are first examined to determine which actions have positive and negative effects on the child's nutritional status. The recoded variable is as follows: no points for no positive action taken (nothing, giving the food to another, etc…), one point if a positive action (encouragement, forcing the food, etc…) is taken. This variable is correlated with HAZ scores at the 0.003 level.

3.7.3 Omitted Variables

Theory and the WHO guidelines both dictate that the child should begin to nurse within half an hour after birth, or soon thereafter. Theory tells us that a child will be healthier if he/she receives colostrum. However, the correlation between HAZ scores and this practice does not conform to these prior expectations. This question has a lot of missing data. For this reason, data from two questions was combined: did the child receive colostrum and when did nursing begin (immediately, within 24 hours, within 48 hours, within 72 hours, later)? For the entire sample, the average HAZ score is –1.6477. Children who begin to nurse immediately after birth have an average HAZ score (-1.662) lower than the cumulative sample mean. However, this action is not statistically significant in this study. Children beginning to nurse within the first day have an average HAZ score of –1.6303. Contrary to theory, those children who began to breastfeed on the second day have an average HAZ score (-1.4162) that is higher than the cumulative sample mean. This group's HAZ scores being lower than the first two groups goes against both theory and the feeding recommendations set out by WHO. When looking at the long-run impacts of delayed breastfeeding, only the group who starts nursing after 72 hours has a negative impact on HAZ scores later in life. Though not statistically significant, this leads to the conclusion that factors other than introduction of breastfeeding are influencing the HAZ scores. These other possible factors may include such factors as birth weight and genetics, thus introducing multicollinearity.

Figure 3.7.1 Feeding Index Variable Descriptives

Variable Name	Variable Description	No. of Obs.	Min	Max	Mean	Relationship to HAZ
Feeding Practices	Number of Positive Feeding Practices Utilized	654	0	5	2.13	Mean HAZ Feed-=0 -2.34 Feed=0.5 -1.44 Feed=1 -2.09 Feed=1.5 -1.56 Feed=2 -1.80 Feed=2.5 -1.38 Feed=3 -1.52 Feed=3.5 -1.26 Feed=4 -1.51 Feed=4.5 -1.34 Feed=5 -1.11 Correlation significant @ 0.01
Diet Diversity	Number of Food Groups Consumed	685	0	3	1.42	Mean HAZ Div-=0 -2.16 Div=0.5 -1.70 Div=1 -1.79 Div=1.5 -1.44 Div=2 -1.78 Div=2.5 -1.32 Div=3 -1.06 Correlation significant @ 0.01
Feeding Refusal Practices	Dummy = 1 if mother employs 'good' feeding refusal practices, 0 otherwise	685	0	1	0.52	Mean HAZ FeedRefusal-=0 -1.92 FeedRefusal=1 -1.50 t-test significant @ <0.00005
Nourished	Dummy = 1 if mother believes the child is well nourished, 0 otherwise (for infants 0-6 months)	810	0	1	0.90	Mean HAZ Nourished=0 -2.30 Nourished=1 -1.64 t-test significant @ <0.00005
Comp_Nurse	Dummy = 1 if Not continuing to nurse after introduction of complimentary feeding, 0 otherwise	631	0	1	0.85	Mean HAZ Comp_Nurse=0 -1.35 Comp_Nurse=1 -1.76 t-test significant @ 0.005
Breastfeeding	Dummy = 1 if child is currently breastfeeding, 0 otherwise	694	0	1	0.72	Mean HAZ Breastfeeding=0 -1.32 Breastfeeding=1 -1.84 t-test significant @ <0.00005
Content	Dummy = 1 if not content with child's nutrition (for children receiving complementary foods), 0 otherwise	723	0	1	0.79	Mean HAZ Content-=0 -2.13 Content=1 -1.59 t-test significant @ <0.00005
prod	Dummy = 1 if child has 2 or more foods in its diet, 0 otherwise	629	0	1	0.85	Mean HAZ prod=0 -2.13 prod=1 -1.62 t-test significant @ 0.002

3.7.4 Factor Analysis

In the rotated component matrix, Component 1 is highly correlated with the following variables: if the child is currently breastfeeding and if the child is continuing to be breastfed while receiving complementary feeding. Thus, this component represents breastfeeding practices. If the child is currently breastfeeding has a factor loading score (correlation coefficient) of 0.902. If the child is continuing to be breastfed while receiving complementary feeding has a factor loading score of 0.908. Component 1 accounts for roughly 35% of the variance in the data.

Component 2 has one variable with a high factor loading score. If the infant is six months and under and the mother is content with the child's nutrition accounts for about 24% of the variance in the data. It has a factor loading score of 0.991.

Component 3 represents one variable, too. The variable is whether or not the mother believes the child is well fed for those children receiving complementary foods. The factor loading score is 0.992 and accounts for 19% of the variance in the data. Therefore, Component 3 represents whether or not the mother believes the child is well fed.

Component 4 also represents one variable. The variable is the diversity of food products consumed by the child. The factor loading score is 0.996. This component accounts for 15% of the data's variance. Thus, Component 4 represents the diversity of food products consumed by the child.

Figure 3.7.2 Feeding Index Matrix: Factor Loadings on Principal Components

	Component			
	1	2	3	4
Believes the child is well fed (for children receiving complementary foods)	-.025	.119	.992	.025
Complimentary feeding and continuing to breastfeed	.908	-.002	-.006	-.041
Dummy = 1 if child currently breastfeeding	.902	-.060	-.030	-.061
Content with nutrition (0-6 months old)	-.045	.991	.119	.038
Food products diversity	-.075	.038	.025	.996

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 4 iterations.

Figure 3.7.3 Feeding Index Variance: Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	1.73	35	35
2	1.22	24	59
3	.94	19	78
4	.76	15	93

Extraction Method: Principal Component Analysis.

3.7.5 Ad Hoc Index Description

The overall feeding index is the sum of scores on three individual variables: when complementary foods began, the diversity of complementary foods, and actions to encourage children to eat complementary foods. The weighting of the indices is modeled after a study done by Arimond and Ruel (2002). As Arimond and Ruel (2002) state, "there is neither standard guidance nor an empirical basis for weighting the various dimensions of infant and child feeding relative to one anther". In this index, a greater weight is given to the dietary diversity variable (0 to 3), while the other two variables are given weights from 0 to 1. This gives the index with a possible range of 0 to 5. The weighting process used during developing these variables is discussed in section 3.7.2 Variable Descriptives.

Though many variables are included in the feeding index, not all variables that were included in the KAP survey are incorporated. Like the care and sanitation indices, some variables yield interesting information, but are not suitable for insertion in the index. Creating this index reconfirms that theory is not always consistent with empirical observation; for example, how soon after birth the child begins to breastfeed and whether or not the child is currently being breastfed was not positively correlated with HAZ scores.

The index, comprised of scores on three individual variables, is tested through a correlation with the HAZ. The feeding index is correlated with HAZ by 0.202 and is significant at the $p<0.00005$ level. The minimum possible and actual value for the index is zero, meaning the child is displaying no feeding practices that are expected to have a positive impact on their HAZ score. The maximum possible and actual value is five. This high score is attained only when practicing the best of feeding practices. The mean value for the feeding index is 2.13; meaning that the average child in this study does not experience good feeding practices.

Figure 3.7.4 Feeding Index Correlation With HAZ

		Index	HAZ
Index	Pearson Correlation	1	.202(**)
	Sig. (2-tailed)	.	.000
	N	664	600

** Correlation is significant at the 0.01 level (2-tailed).

CHAPTER 4: TESTING THE INDICES

4.1 Original Results without the Indices

The original model upon which this report is based is taken from the study, "Identifying Policy Relevant Variables for Reducing Childhood Malnutrition in Rural Mali" (Kelly et al. 2004). Both studies use some KAP variables in a regression analysis. The aim of the Kelly et al. regression is to determine which aspects have the largest effect on a child's nutritional status. The child's height-for-age Z-score (HAZ) is used as the proxy for nutritional status.

The original Kelly et al. model and the modified models use the OLS method for regression estimation through reduced form models. Though this is the standard approach in research, this method does not eliminate issues of endogeneity[9]. A more thorough method of regressing models is the 2-stage method, which allows the researcher to obtain unbiased coefficient estimates if endogeneity is a problem. Since the 2-stage method is not employed in this study, potential problems of endogeneity should be noted when examining the results, paying specific attention to the relationship between feeding practices and the HAZ score.

<u>Original Model:</u>

<u>The original model developed by Kelly et al. had the following structure:</u>

HAZ=a+b1(sex)+b2(age)+b3(diet diversity)+b4(mother schooling)+b5(mother age)+b6(mother age squared)+b7(mother height)+b8(prenatal visits)+b9(father age)+b10(father height)+b11(chief schooling)+b12(chief literacy training)+b13(ln kcal cereal consumption)+b14(ln expenditure)+b15(clinic staffed)+b17(distance to clinic)+b18(Kolondieba)+b19(Koutiala)+b20(Mancina)+b21(Niono)+b22(Bandiagara) + ε

The results from the original model (Kelly et al., 2004) are found below in Table 4.1.1. The adjusted R squared is 0.21 and $p < 0.0005$. These statistics show overall model is highly significant, but only explains 21% of the variance found in the data.

This estimation indicates that several of the independent variables have a statistically significant effect on HAZ scores. Child feeding practices, distance to health centers, and household income are all statistically significant. The mother's and child's ages are also significant determinants in HAZ scores. The information derived from these results can be used in nutrition intervention programs. For example, if use of or distance health centers has a significantly positive affect on a child's HAZ score, the government will know to build health centers that are centrally located to the villages. Parents' standardized heights are two more significant variables. This points toward the possibility that the "HAZ score may be determined by either the genetic traits of their parents or socio-economic and environmental factors such as persistent poverty that contributed to parental stunting" (Kelly et al., 2004). Kebede (2003) concluded in his Ethiopian study "the height of parents is significant in all specifications." The last set of significant variables is the zone in which the child lives. The HAZ score will be either positively or negatively influenced depending upon in which agricultural production zone

[9] Income was tested for endogeneity.

the child lives. Here, Koro, which is in the millet/sorghum zone, is the omitted value among the region dummy variables. Hence, the region dummies measure deviation from scores in Koro. Bandiagara, in the millet/sorghum zone, has the highest scores and Kolondieba, in the cotton zone, has the lowest scores.

Table 4.1.1 Summary of Linear and Logistic Model Results (Original Model)

Dependent Variable = Child's HAZ Score		
	Linear Model	
Variables	Coef.	Sig
Male child	-0.13	
Child 0-6 mos.	0.57	*
Child 7-12 mos.	0.32	**
Child's diet diverse	0.51	***
Mother's schooling (yrs)	-0.04	
Mother's age (yrs)	0.08	**
Mother's age squared	0.00	***
Mother's height (standardized)	0.22	***
One or more prenatal visit=1	0.29	**
Father's age (yrs)	0.01	*
Father's height (standardized)	0.21	***
Schooling PCU head (yrs)	-0.07	*
Literacy training PCU head (yrs)	-0.05	
Ln kcal cereal consumption/ae	-0.13	
Ln expenditure/ae	0.19	*
Clinic fully staffed = 1	0.30	**
Distance to clinic (km)	-0.01	**
New cotton zone (Kolondieba) = 1	-0.75	***
Old cotton (Koutiala) =1	-0.59	**
Rice/millet zone (Macina) = 1	-0.15	
Rice zone (Niono) = 1	0.14	
Mil/sor/onion zone (Bandiagara) =1	0.53	**
Constant	-4.65	***
Observations	457	
Prob F	0.00	
Adj. R sq	0.21	
Distribution of residuals	Normal	
Note: Models run using Huber/White/sandwich estimator of variance		
* 0.10 level; ** 0.05 level; *** 0.01 level or better		

Source: Kelly et al., 2004

4.2 Results With The Indices

The structure of the model with indices (factor analysis feeding index is the following):

HAZ=a+b1 (Sex)+b2 (Age)+b3 (feeding index)+b5 (mother schooling)+b6 (mother age)_b7 (mother age squared)+b8 (mother height)+b9 (prenatal care index)+b10 (father age)+b11 (father height)+b12 (chief literacy training)+b13 (ln kcal cereal consumption)+b14 (ln expenditure)+b15 (diarrhea care index)+b17 (mother health care index)+b18 (Kolondieba)+b19 (Koutiala)+b20 (Mancina)+b21 (Niono)+ b22 (Bandiagara)+b23 (water sanitation index)+b24 (disease avoidance sanitation index)+b25 (mother hand washing sanitation index)+b26(child hand washing sanitation index) + ε

The structure of the model with the ad hoc feeding index is the following:

HAZ=a+b1 (Sex)+b2 (Age)+b3 (feeding index)+b4 (currently breastfeeding)+b5 (mother schooling)+b6 (mother age)_b7 (mother age squared)+b8 (mother height)+b9 (prenatal care index)+b10 (father age)+b11 (father height)+b12 (chief literacy training)+b13 (ln kcal cereal consumption)+b14 (ln expenditure)+b15 (diarrhea care index)+b17 (mother health care index)+b18 (Kolondieba)+b19 (Koutiala)+b20 (Mancina)+b21 (Niono)+ b22 (Bandiagara)+b23 (water sanitation index)+b24 (disease avoidance sanitation index)+b25 (mother hand washing sanitation index)+b26 (child hand washing sanitation index) + ε

4.2.1 Model Using Feeding Factor Analysis

The summary of the model that includes the indices with factor analysis is found in Tables 4.2.1 and Annex 1. The aim of using indices is to include more information than the original model, thereby increasing the explanatory power of the model. The adjusted

R squared value for this new model is 0.18 and the probability of the F statistic is <0.00005.

These results show that the significant contributing factors influencing a child's nutritional status are breastfeeding practices (feeding index), where the mother gave birth to the child (a care index), health center utilization (a care index), care of the child concerning diarrhea episodes (a care index), mother's standardized height, and father's standardized height, and the sex of the child. Four of the seven significant determinants are indices; thereby encompassing many more variables than are directly represented in the regression.

4.2.1 Comparison of OLS results using proxy variables vs. indices and factor scores (factor analysis for feeding index)

	Model using proxies		Model using Indices	
Variables	Coefficient	Signif.	Coefficient	Signif.
Male Child	-0.13		-0.43	***
Child 0-6 Months	0.57*		0.54	
Child 7-12 Months	0.32**		0.37	
Dietary diversity dummy	0.51***			
Breastfeeding practices (Feeding Index)			-0.22	**
Content w/ child's nutrition, comp. foods (Feeding Index)			0.05	
Food products consumed (Feeding Index)			0.02	
Believe child is well fed, 0-6 months (Feeding Index)			0.11	
Mother's Schooling (yrs)	-0.04**		0.01	
Mother's Age (yrs)	0.08***		0.06	
Mother's Age Squared	0.00***		-0.00	
Mother's Height (standardized)	0.22**		0.22	***
Prenatal care dummy	0.29	**		
Prenatal & Delivery Care factor score (Care Index)			0.15	*
Father's Age (yrs)	0.01*		0.01	
Father's Height (standardized)	0.21***		0.29	***
Formal schooling PCU head (yrs)	-0.07*		-0.7	
Literacy Training PCU Head (yrs)	-0.05		-0.12	
Ln kcal cereal consumption/ae	-0.13		-0.31	
Ln expenditure/ae	0.19*		0.18	
Health center distance (km)	-0.01	**		
Health center fully staffed (dummy = 1)	0.30	**		
Child Care For Diarrhea factor score (Care Index)			0.21	***
Mother's General Health Care factor score (Care Index)			.17	**
Kolondieba (new cotton zone) dummy	-0.75***		-0.51	*
Koutiala (old cotton zone) dummy	-0.59**		-0.48	
Macina (rice/millet zone) dummy	-0.15		-0.27	
Niono (rice zone) dummy	0.14		0.22	
Bandiagara (millet/sorghum zone) dummy	0.53**		0.36	
Potable Water Source (Sanitation Index)			0.05	
Avoidance of Fecal/Mosquito Borne Diseases (Sanitation Index)			0.06	
Mother Hand Washing Practices (Sanitation Index)			0.02	
Child Hand Washing Practices (Sanitation Index)			-1.75	
Constant	-4.65***		-1.93	
Observations	457		281	
Prob F	0.00		0.00	
Adj. R sq	0.21		0.18	
Distribution of residuals				
* 0.10 level; ** 0.05 level; *** 0.01 level or better				

4.2.2 Ad Hoc Feeding Index

The summary of the models that include the indices with factor analysis and the ad hoc feeding index is found in Tables 4.2.3 and Annex 2. The adjusted R squared value for this new model is 0.183 and the probability of the F statistic is <0.00005.

These results show that the largest contributing factors influencing a child's nutritional status are the feeding index, where the mother gave birth to the child (a care index), health center utilization (a care index), care of the child concerning diarrhea episodes (a care index), mother's standardized height, and father's standardized height, being between seven and twelve months, and sex of the child. Four of the eight significant determinants are indices; thereby encompassing many more variables than are directly represented in the regression.

Table 4.2.3 Comparison of OLS results using proxy variables vs. indices and factor scores (ad hoc feeding index)

	Model using proxies		Model using Indices	
Variables	Coefficient	Signif.	Coefficient	Signif.
Male Child	-0.13		-0.334	**
Child 0-6 Months	0.57	*	0.55	
Child 7-12 Months	0.32	**	0.323	*
Dietary diversity dummy	0.51	***		
Feeding Index			0.133	*
Currently Breastfeeding			-0.37	*
Mother's Schooling (yrs)	-0.04	**	-0.009	
Mother's Age (yrs)	0.08	***	0.042	
Mother's Age Squared	0.00	***	-0.001	
Mother's Height (standardized)	0.22	**	0.216	***
Prenatal care dummy	0.29	**		
Prenatal & Delivery Care factor score (Care Index)			0.17	**
Father's Age (yrs)	0.01	*	0.011	
Father's Height (standardized)	0.21	***	0.262	***
Formal schooling PCU head (yrs)	-0.07	*		
Literacy Training PCU Head (yrs)	-0.05		-0.1	
Ln kcal cereal consumption/ae	-0.13		-0.375	*
Ln expenditure/ae	0.19	*	0.195	
Health center distance (km)	-0.01	**		
Health center fully staffed (dummy = 1)	0.30	**		
Child Care For Diarrhea factor score (Care Index)			0.203	***
Mother's General Health Care factor score (Care Index)			0.154	*
Kolondieba (new cotton zone) dummy	-0.75	***	-0.721	*
Koutiala (old cotton zone) dummy	-0.59	**	-0.525	*
Macina (rice/millet zone) dummy	-0.15		-0.278	
Niono (rice zone) dummy	0.14		0.074	
Bandiagara (millet/sorghum zone) dummy	0.53	**	0.215	
Potable Water Source (Sanitation Index)			0.062	
Avoidance of Fecal/Mosquito Borne Diseases (Sanitation Index)			-0.037	
Mother Hand Washing Practices (Sanitation Index)			-0.02	
Child Hand Washing Practices (Sanitation Index)			-0.62	
Constant	-4.65	***	-1.258	
Observations	457		276	
Prob F	0.00		0.00	
Adj. R sq	0.21		0.18	
Distribution of residuals				
* 0.10 level; ** 0.05 level; *** 0.01 level or better				

4.3 Interpretation of Results

Both models with indices yield the same significant variables. However, it must be noted that since the feeding index is divided into factors in the factor analysis regression, this sheds light on a more specific level than the ad hoc index. The ad hoc feeding index is coming out as significant due to an underlying factor. The factor analysis derived feeding index brings this underlying factor to light; breastfeeding practices.

The significant variables that come forth in the modified model (with indices) shed light on changes that can be made by the Malian government. The feeding index, two care indices factors, and kcals of cereal consumed per adult equivalent in the child's household are all statistically significant. The feeding index is significant at the 10% level, the diarrhea care index is significant at the 1% level, the mother's general care index at 10%, and the kcals of cereal consumed per adult equivalent is significant at the 10% level[10]. These specific variables can be used to influence new policy reforms and programs. Child's age, sex, height of mother and father are also significant variables. Depending upon in which village the child lives, their nutritional status will be significantly influenced. This model reaches the same conclusions as that of the Kelly et al. paper, in that genetic socio-economic factors influence a child's HAZ score.

An important difference between the two equations is the number of observations. Kelly et al.'s model uses 457 observations while the modified model uses just 276

[10] Significance levels taken from the regression including the ad hoc feeding index

observations[11]. The same data set is used for both models, but the original model has access to more observations. This is because Kelly et al. used different and fewer variables than this study. This study lost many cases because when there was a variable with a missing observation, the whole case was not used, thereby omitting the other variables as well. Slight differences, such as the population dispersion between zones, exist between variables common to both models, due to the data set size as well as interrelated independent variables. The most notable difference is in the Niono region (rice zone). In the original model, fifteen percent of the children live in the Niono region. The new model has seven percent of the children living there. Therefore, the original model is more heavily weighted with Niono region knowledge, attitudes, and practices.

Though a model with a large number of observations will, ceteris paribus, have a larger R-squared value, the higher R^2 does not necessarily signify that the model is an accurate reflection of the underlying relationships. Therefore, another regression is also included in this study. The 276 observations are regressed on the original model[12]. This provides more insight into the models with indices. The R^2 value in this new run is the same as that of the models with the indices (0.18). This, too, points towards the models with indices as being an improvement over the original. When both models are run with the same observations, the R^2 is the same for both. The higher R^2 in the original model is higher due to the sample size.

[11] Ceteris paribus, the model with more observations will have a higher R-squared value due to more points to aid in omitting 'noise' in the relationships being estimated.
[12] See Annex 3 for the results of the original model when run with 276 observations.

CHAPTER 5: CONCLUSIONS

5.1 Summary of the Approach

The broad objective of this study is to identify the contributions that including indices in a model can make, versus just including the individual variables. The paper builds on work already done on this topic. The implementation of this research project will aid in formulating more precise policy recommendations. Through the creation of these indices, we hope to incorporate a variety of factors that would otherwise be extremely difficult to capture into the HAZ model that examines the determinants of child nutritional status.

This report investigates which determinants have the greatest effect on the HAZ scores, an indicator of the long-term nutritional status of Malian children under five years of age. Ordinary least squares (OLS) regression analysis was employed in analyzing a household level survey. The variables used in the indices were transformed into binary variables. The one exception is the mother's hand washing practices, which is a truncated continuous variable. From these variables, indices were created using factor analysis and, in one instance, an ad hoc method. These indices, exogenous variables, were regressed on the child's most recent HAZ score.

5.2 Performance of the Indices

Explanatory power of the individual indices varies. The regression model incorporating the ad hoc feeding index includes eight independent variables, while the model that used factor analysis for the feeding indices includes eleven independent variables. All indices are derived from three index categories: childcare, sanitation, and feeding. Coefficients of four of the indices are not significant in the regression. Four of the indices are significant and deal with care and feeding. None of the sanitation indices are significant. The care indices and breastfeeding index are indices that have standardized beta coefficients above 0.1 (see Table 5.2.1). These are the indices with the highest standardized coefficients. Other variables in the model have higher standardized coefficients, but are not indices: mother's age (0.35), mother's height (0.16), and father's height (0.22). The standardized coefficients allow for equal comparisons of variables, despite the different units of measure used for specific variables. The larger the standardized coefficient, the larger the impact that variable will have on the child's HAZ score. Though contradictory to our hypothesis, the index for breastfeeding practices has a significant, negative beta coefficient: the higher that index for an individual, the lower their HAZ score will be.

The two models created in this study that involve two different ways of creating indices (factor analysis vs. an ad-hoc approach) yield roughly the same results. The model with the factor–analysis-derived feeding index has one advantage over the ad hoc derived feeding index. The factor-analysis-derived feeding index brings forth the underlying significant variable from the ad hoc feeding index, breastfeeding practices.

From a policy perspective, knowing that the general care and diarrhea care variables are significant is an improvement over the original model developed by Kelly et al. that does not include indices. The government can support programs that address knowledge, while it can't do anything about the mother's age, other than encourage women to have children later in life, which is not a realistic scenario in the short-run in rural Mali, although it might be able to affect this variable over the long-run through education programs. Therefore, explanatory power of the second model is higher because in the first model we can only guess for what 'mothers age' serves as a proxy.

Though the original model has a larger number of observations and will, ceteris paribus, have a larger R-squared value, the higher R^2 does not necessarily signify that the model is an accurate reflection of the underlying relationships. For example, in the model with the indices, it is found that the care index for diarrhea is highly significant. However, the mother's age is no longer significant, as it was in the original model. If mother's age is highly correlated with knowledge of how to provide general care and care for diarrhea, then the first model is reflecting the care variable via the age variable.

Table 5.2.1 Standardized Beta Coefficients (All Indices are Factor Analysis Derived)

	Model using Indices	
Variables	Std. Coefficient	Signif.
Male Child	-0.16	***
Child 0-6 Months	0.08	
Child 7-12 Months	0.13	
Dietary diversity dummy		
Breastfeeding practices (Feeding Index)	-0.16	**
Content w/ child's nutrition (Feeding Index)	0.04	
Food products consumed (Feeding Index)	0.01	
Believe child is well fed (Feeding Index)	0.07	
Mother's Schooling (yrs)	0.01	
Mother's Age (yrs)	0.35	
Mother's Age Squared	-0.41	
Mother's Height (standardized)	0.16	***
Prenatal care dummy		
Prenatal & Delivery Care factor score (Care Index)	0.11	*
Father's Age (yrs)	0.06	
Father's Height (standardized)	0.22	***
Formal schooling PCU head (yrs)	-0.07	
Literacy Training PCU Head (yrs)	-0.09	
Ln kcal cereal consumption/ae	-0.11	
Ln expenditure/ae	0.09	
Health center distance (km)		
Health center fully staffed (dummy = 1)		
Child Care For Diarrhea factor score (Care Index)	0.16	***
Mother's General Health Care factor score (Care Index)	0.14	**
Kolondieba (new cotton zone) dummy	-0.16	*
Koutiala (old cotton zone) dummy	-0.15	
Macina (rice/millet zone) dummy	-0.8	
Niono (rice zone) dummy	0.04	
Bandiagara (millet/sorghum zone) dummy	0.09	
Potable Water Source (Sanitation Index)	0.04	
Avoidance of Fecal/Mosquito Borne Diseases (Sanitation Index)	0.04	
Mother Hand Washing Practices (Sanitation Index)	0.08	
Child Hand Washing Practices (Sanitation Index)	-0.05	
Constant	X	
Observations	281	
Prob F	0.00	
Adj. R sq	0.18	
Distribution of residuals		
* 0.10 level; ** 0.05 level; *** 0.01 level or better		

5.3 Limitations of the Study

One of the limitations to any approach based on indices that incorporate a large number of variables is that one needs valid responses on all the questions that go into the index in order to calculate the index. In this study, the number of observations on certain key variables that we would have liked to include in the index was limited, precluding their inclusion in the index. Due to this lack of response, this study has a somewhat limited view of Malian life and cannot generalize to the degree that the author was hoping to at the start of the project. Another problem related to the lack of response is endogeneity. Non-observed variables could be influencing the results of this study through the error term and/or included variables could be influencing the results through simultaneous determination.

5.4 Policy Implications

Mali is a country with a high rate of infant malnutrition and morbidity. Understanding better the determinants of child HAZ scores will help develop useful suggestions for policymakers. With the determinants identified and the magnitude of the influence on HAZ score identified, policy makers will have a platform to make decisions to change the nutritional status of the country of the better. They will be able to determine what characteristics most influence the overall nutritional well-being of the population.

The variables with the highest degree of significance show where improvement will have the greatest impact on children's nutritional status. This study's models show that actions affecting the care of the child play the biggest role. All three of the significant variables are indices from the care index: prenatal and delivery care, diarrhea care, and mother's general care. This means that making health centers more accessible would have a large impact on childhood malnutrition.

Implementation of programs to educate mothers on the benefits of visiting the health center, not only for the child's health, but also their own, could have a significantly large impact on the child's HAZ score. The variables encompassed in this explanation are both significant. The care index concerning prenatal care and delivery is significant at the five percent level and has a standardized beta coefficient of 0.13. The care index representing the mother's general care is significant at the ten percent level and has a standardized beta coefficient of 0.12. Other programs involving the health center could provide for adequate numbers of staffing as well as encouraging birthing at the health centers.

The index dealing with actions taken for episodes of diarrhea is significant at the five percent level in the model with a standardized beta coefficient of 0.16. Education programs directed towards urging caregivers to bring children to the health center for treatment of diarrhea could reduce the instances of childhood malnutrition.

5.5 Suggestions for Further Research

This study shows statistically significant indices are productive components when included in a regression model. Using all of the indices in the model points to areas in which further research is needed. The factors used from the sanitation index do not yield significant results. Developing significant indices for implementation for sanitation practices is one area of research that can be pursued. Gathering accurate information of many observations on these issues is a necessity for further research.

Annexes

Annex 1: Comparison of descriptive statistic results using proxy variables vs. indices (factor analysis feeding index) and factor scores

	Model using proxies			Model using Indices		
Variables	mean	min/max	SD	mean	min/max	SD
HAZ score	-1.60	-4.98/2.58	1.30	-1.71	-4.98/2.27	1.31
Male Child	.51	0/1	.50	0.50	0/1	0.50
Child 0-6 Months	.05	0/1	.23	0.05	0/1	0.20
Child 7-12 Months	.23	0/1	.42	0.23	0/1	0.44
Dietary diversity dummy	.85	0/1	.36			
Breastfeeding practices (Feeding Index)				0	-2.37/0.61	1
Content w/ child's nutrition (Feeding Index)				0	-2.19/0.98	1
Food products consumed (Feeding Index)				0	-2.52/0.60	1
Believe child is well fed (Feeding Index)				0	-3.09/0.65	1
Mother's Schooling (yrs)	.33	0/9	1.23	0.32	0/9	1.39
Mother's Age (yrs)	28.35	15/64	7.34	28.35	15/64	7.65
Mother's Age Squared	858.18	225/4096	462.76	857.85	225/4096	484.83
Mother's Height (standardized)	.06	-4.97/3.32	1.01	0.05	-4.97/3.32	0.94
Prenatal care dummy	.67	0/1	.47			
Prenatal & Delivery Care factor score (Care Index)				0.00	-1.58/1.79	1.00
Father's Age (yrs)	39.56	20/76	10.79	39.50	20/76	10.76
Father's Height (standardized)	0	-3.11/2.86	1.00	0.01	-3.11/2.86	1.00
Formal schooling PCU head (yrs)	.33	0/8	1.20	0.16	0/8	1.19
Literacy Training PCU Head (yrs)	.14	0/10	.77	0.33	0/10	0.80
Ln kcal cereal consumption/ae	7.59	6.02/8.70	.46	7.59	6.02/8.70	0.46
Ln expenditure/ae	9.95	7.28/11.99	.69	9.96	7.28/11.99	0.69
Health center distance (km)	11.27	0/80.50	11.75			
Health center fully staffed (dummy = 1)	.69	0/1	.46			
Child Care For Diarrhea factor score (Care Index)				0.00	-1.62/0.97	1.00
Mother's General Health Care factor score (Care Index)				0.00	-1.80/1.77	1.00
Kolondieba (new cotton zone) dummy	.19	0/1	.39	0.19	0/1	0.39
Koutialla (old cotton zone) dummy	.16	0/1	.37	0.16	0/1	0.37
Macina (rice/millet zone) dummy	.23	0/1	.42	0.23	0/1	0.42
Niono (rice zone) dummy	.15	0/1	.36	0.14	0/1	0.36
Bandiagara (millet/sorghum zone) dummy	.12	0/1	.33	0.12	0/1	0.32
Potable Water Source (Sanitation Index)				0.00	-1.23/1.41	1.00
Avoidance Fecal/Mosquito Borne Diseases (Sanitation Index)				0.00	-1.02/2.04	1.00
Mother Hand Washing Practices (Sanitation Index)				0.00	-0.60/3.99	1.00
Child Hand Washing Practices (Sanitation Index)				0.00	-0.21/26.38	1.00
Number of observations	457			276		

Annex 2: Comparison of descriptive statistic results using proxy variables vs. indices (ad hoc feeding index) and factor scores

	Model using proxies			Model using Indices		
Variables	mean	min/max	SD	mean	min/max	SD
HAZ score	-1.60	-4.98/2.58	1.30	-1.67	-4.98/2.27	1.30
Male Child	.51	0/1	.50	.53	0/1	.50
Child 0-6 Months	.05	0/1	.23	.04	0/1	.20
Child 7-12 Months	.23	0/1	.42	.44	0/1	.26
Dietary diversity dummy	.85	0/1	.36			
Feeding Index				2.11	0/5	1/13
Currently Breastfeeding				.75	0/1	.43
Mother's Schooling (yrs)	.33	0/9	1.23	.39	0/9	1.46
Mother's Age (yrs)	28.35	15/64	7.34	28.62	16/74	7.59
Mother's Age Squared	858.18	225/4096	462.76	876.41	256/4096	483.16
Mother's Height (standardized)	.06	-4.97/3.32	1.01	-.06	-2.91/2.75	.94
Prenatal care dummy	.67	0/1	.47			
Prenatal & Delivery Care factor score (Care Index)				.06	-1.58/1.79	.98
Father's Age (yrs)	39.56	20/76	10.79	39.56	20/76	10.86
Father's Height (standardized)	0	-3.11/2.86	1.00	.01	-2.40/2.59	.99
Formal schooling PCU head (yrs)	.33	0/8	1.20	.30	0/8	1.18
Literacy Training PCU Head (yrs)	.14	0/10	.77	.17	0/10	.89
Ln kcal cereal consumption/ae	7.59	6.02/8.70	.46	7.54	6.02/8.56	.45
Ln expenditure/ae	9.95	7.28/11.99	.69	9.83	7.28/11.47	.68
Health center distance (km)	11.27	0/80.50	11.75			
Health center fully staffed (dummy = 1)	.69	0/1	.46			
Child Care For Diarrhea factor score (Care Index)				-.01	-1.62/.97	.99
Mother's General Health Care factor score (Care Index)				.05	-1.80/1.77	.99
Kolondieba (new cotton zone) dummy	.19	0/1	.39	.22	0/1	.42
Koutialla (old cotton zone) dummy	.16	0/1	.37	.21	0/1	.41
Macina (rice/millet zone) dummy	.23	0/1	.42	.21	0/1	.41
Niono (rice zone) dummy	.15	0/1	.36	.07	0/1	.25
Bandiagara (millet/sorghum zone) dummy	.12	0/1	.33	.14	0/1	.35
Potable Water Source (Sanitation Index)				-.08	-1.14/1.41	.98
Avoidance Fecal/Mosquito Borne Diseases (Sanitation Index)				-.04	-1.02/2.04	.98
Mother Hand Washing Practices (Sanitation Index)				-.03	-.60/3.89	.95
Child Hand Washing Practices (Sanitation Index)				-.04	-.21/.12	.04
Number of observations	457			276		

Annex 3: Running the Original Model with the Observations used in the Models with Indices (Summary of Linear and Logistic Model Results)

	Original Model Observations		Original Model with Index Observations	
Variables	Coefficient	Signif.	Coefficient	Signif.
Male child	-0.13		-0.34	**
Child 0-6 mos.	0.57	*	0.43	
Child 7-12 mos.	0.32	**	0.17	
Child's diet diverse	0.51	***	0.29	
Mother's schooling (yrs)	-0.04		-0.02	
Mother's age (yrs)	0.08	**	0.06	
Mother's age squared	0.00	***	0.00	
Mother's height (standardized)	0.22	***	0.23	***
One or more prenatal visit=1	0.29	**	0.61	***
Father's age (yrs)	0.01	*	0.01	
Father's height (standardized)	0.21	***	0.29	***
Schooling PCU head (yrs)	-0.07	*	-0.08	
Literacy training PCU head (yrs)	-0.05		-0.09	
Ln kcal cereal consumption/ae	-0.13		-0.21	
Ln expenditure/ae	0.19	*	0.27	*
Clinic fully staffed = 1	0.30	**	0.26	
Distance to clinic (km)	-0.01	**	-0.02	
New cotton zone (Kolondieba) = 1	-0.75	***	-0.90	***
Old cotton (Koutiala) =1	-0.59	**	-0.60	*
Rice/millet zone (Macina) = 1	-0.15		-0.37	
Rice zone (Niono) = 1	0.14		0.05	
Mil/sor/onion zone (Bandiagara) =1	0.53	**	0.52	*
Constant	-4.65	***	-4.08	*
Observations	457		276	
Prob F	0.00		0.00	
Adj. R sq	0.21		0.18	
Distribution of residuals	Normal		Normal	
* 0.10 level; ** 0.05 level; *** 0.01 level or better				

References

Adair, L., and D. Guilkey. 1997. Age-specific determinants of stunting in Filipino children. *Journal of Nutrition* 127: 314-320.

Adams, A., and S. Madhavan and D. Simon. 2002. Women's social networks and child survival in Mali. *Social Sciences and Medicine* 54:165-178.

Arimond, M., and M. Ruel. December 2002. Progress in developing an infant and child feeding index: An example using the Ethiopia Demographic and Health Survey 2000. FCND discussion paper 143. Food Consumption and Nutrition Division, International Food Policy Research Institute.

Bhandari, N., and R. Bahl, S. Taneja, M. de Onis, and M.K. Bhan. 2002. Growth performance of affluent Indian children is similar to that in developed countries. *Bulletin of the World Health Organization* 80(3):189-195.

Bingen, R., and D. Robinson and J. Staatz. 2000. Democracy and Development in Mali. East Lansing, Michigan State University Press.

Cebu Study Team. 1992. A child health production function estimated from longitudinal data. *Journal of Development Economics* 38:323-352.

Cogill, B. 2003. Anthropometric Indicators Measurement Guide. Section 6: Comparison of Anthropometric Data to Reference Standards. *Food and Nutrition Technical Assistance.* Rev. ed. 39-42.

Darlington, R.B., and S. Weinberg and H. Walberg. 1973. Canonical variate analysis and related techniques. *Review of Educational Research.* Cornell University.

Niama Nango Dembélé and John M. Staatz. 2002. "The Impact on Market Reform on Agricultural Transformation in Mali." In Perspectives on Agricultural Transformation: A View from Africa, edited by T.S. Jayne, Isaac J. Minde,and Gem Argwings-Kodhek, pp. 163-89. Jericho, NY: Nova Science Publishers.

Faruqee, R. February 1979. Sources of fertility decline: Factor analysis of inter-country data. Staff working paper 318, World Bank.

Kebede, B. 2003. Genetic endowments, parental and child health in rural Ethiopia. Center for the Study of African Economies working paper 2003-10, University of Bath and Oxford University, England.

Kelly, V., and J. Tefft, J. Oehmke, and J. Staatz. 2004. Identifying policy relevant variables for reducing childhood malnutrition in rural Mali. Staff paper 04-28, Michigan State Univ, East Lansing.

Labaste, Patrick. 1996. Mali: Intensification of Irrigated Agriculture at the Office du Niger. Unpublished Working Paper, World Bank, Washington.

de Onis, M., and C. Monteiro, J. Akre, and G. Clugston. 1993. The worldwide magnitude of protein-energy malnutrition: An overview from the WHO global database on child growth. *Bulletin of the World Health Organization* 71(6):703-712.

Paknawin-Mock, J., and L. Jarvis, AB. Jahari, MA. Husaini, and E. Pollitt. 2000. Community-level determinants of child growth in an Indonesian tea plantation. *European Journal of Clinical Nutrition* 54:28-42.

Penders, C.L. 1999. Examining the links between agriculture and child malnutrition in Mali. Master's thesis, Michigan State Univ, East Lansing.

Penders, C., and J. Staatz, and J. Tefft. 2000. Agricultural Development and Child Nutrition: What Do We Know?. Policy synthesis paper 52, Michigan State Univ, East Lansing.

Ruel, M., and C. Levin, M. Armar-Klemesu, D. Maxwell, and S. Morris. April 1999. Good care practices can mitigate the negative effects of poverty and low maternal schooling on children's nutritional status: Evidence from Accra. International Food Policy Research Institute: Food Consumption and Nutrition Division. Washington, D.C.

Smith, L., and L. Haddad. 2000. Overcoming child malnutrition in developing countries: Past achievements and future choices. Discussion paper 30, International Food Policy Research Institute. Washington, D.C.

Tefft, J. and V. Kelly. 2002. "Understanding and Reducing Child Malnutrition in Mali." Interim Research Findings for the Project on Linkages Between Child Nutrition and Agricultural Growth (LICNAG). Michigan State University, East Lansing.

Tefft, J. and V. Kelly and J. Staatz. 2003. Linkages between Child Nutrition and Agricultural Growth: Insights from Mali. Michigan State University, Food Security III Cooperative Agreement. http://www.aec.msu.edu/agecon/fs2/mali_nut/index.htm

Tefft, J., and V. Kelly, V. Wise, and J. Staatz, in collaboration with LICNAG Research Team. 2003. Linkages between child nutrition and agricultural growth in Mali: A summary of preliminary findings. Policy synthesis paper 64, Michigan State Univ., East Lansing.

Tefft, J., and J. Staatz and C. Penders. 1999. An action plan to strengthen linkages between agricultural productivity growth and improved childhood nutrition. Policy synthesis paper 48, Michigan State Univ, East Lansing.

Tefft, J., and F. Ouattara, C. Penders, M. Yade, D. Keita, V. Kelly and J. Staatz. 2001. Preliminary Findings on the Linkages Between Agricultural Productivity Growth and Child Nutritional Status in Mali. Policy synthesis paper 57, Michigan State Univ, East Lansing.

Tyner, W., and A. Berry, D. Brown, S. Gavian, D. Humpal, V. Kelly, J. Mellor, S. Piccicuto, J. Staatz, and J. Tefft. March 2002. Mali Agricultural Sector Assessment: Volume 2, In-depth Analyses and Supporting Materials. Agricultural Policy Development Project. Abt Associates Inc., Bethseda, MD.

UNICEF. 1996. Food, health and care: The UNICEF vision for a world free from hunger and malnutrition. New York: United Nations Children's Education Fund.

Wise, V., and J. Tefft, V. Kelly, and J. Staatz, in collaboration with the LICNAG Research Team. October 2002. Knowledge, attitudes and practices on child feeding and care: Preliminary insights from the Project on Linkages between Child Nutrition and Agricultural Growth. Policy synthesis paper 61, Michigan State Univ, East Lansing.

Wooldridge, J.M. 2000. *Introductory econometrics: A modern approach.* South-Western College Publishing, Cincinnati, OH.

World Health Organization, Poverty and Health – Quantitative Techniques for Health Equity Analysis: Technical Note 2. World Health Organization http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTHEALTHNUTRITIONANDPOPULATION/EXTPAH/0,,contentMDK:20216933~menuPK:460204~pagePK:148956~piPK:216618~theSitePK:400476,00.html.