Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)

Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

Nr. 152

A Multilevel Approach to Explain Child Mortality and Undernutrition in South Asia and Sub-Saharan Africa

Kenneth Harttgen, Mark Misselhorn

September 2006

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173
e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

A Multilevel Approach to Explain Child Mortality and Undernutrition in South Asia and Sub-Saharan Africa

Kenneth Harttgen* and Mark Misselhorn†

September 19, 2006

Abstract

While undernutrition among children is very pervasive both in Sub-Saharan Africa and South Asia, child mortality is rather low in South Asia. In contrast to that Sub-Saharan African countries suffer by far the worst from high rates of child mortality. This different pattern of child mortality and undernutrition in both regions is well known, but approaches using aggregated macro data have not been able to explain it appropriately. In this paper we analyze the determinants of child mortality as well as child undernutrition based on DHS data sets for a sample of five developing countries in South Asia and Sub-Saharan Africa. We investigate the effects of individual, household and cluster socioeconomic characteristics using a multilevel model approach and examine their respective influences on both phenomena. We find that the determinants of child mortality and undernutrition differ significantly from each other. Access to health infrastructure is more important for child mortality, whereas the individual characteristics like wealth and educational and nutritional characteristics of mothers play a larger role for anthropometric shortfalls. Although very similar patterns in the determinants of each phenomenon are discernable between countries, there are large differences in the magnitude of the coefficients. Besides regressions using a combined data set of all six countries show, that there are still significant differences between the two regions although taking account of a large set of covariates.

***JEL* Classification:** C40, I12, I31, I32, O57.

Key words: Child mortality, child undernutrition, multilevel modelling.

*University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany, email: k.harttgen@wiwi.uni-goettingen.de.

†University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, email: mark.misselhorn@wiwi.uni-goettingen.de.

1 Introduction

Despite the overall decline in the prevalence of undernutrition and child mortality in developing countries, both phenomena are still at unacceptably high levels and therefore remain big challenges in the fight against lacking capabilities and reaching the MDGs. Concerning the children's anthropometric failure, the WHO (2002) estimated that almost 27 percent (168 million) of children under five years of age are underweight. And looking at the threat of child mortality, nearly 11 million children died in the year 2003 before reaching the age of five. Around 98 percent of the deaths occur in developing countries (UN 2005). Several papers have studied the socioeconomic determinants of child mortality and undernutrition. Examples for empirical studies of child mortality are Subbaro and Rany (1995), Pritchett and Summers (1996), Ssewanyana and Younger (2004), and for undernutrition Gillespie, Mason and Martorell (1996), Osmani (1997) and more recently by Smith and Haddad (2000). As stated in numerous studies in this field, one of the major causes of child mortality is undernutrition itself. Most studies cite this result by referring to a study by Pelletier et al. (1995) which finds that more than 50 percent of child mortality is attributable to mild, moderate and severe undernutrition. In addition, a study of Pelletier et al. (2002) measures the effect of malnutrition on changes in child mortality for 59 developing countries using aggregate longitudinal data from 1966 to 1996 finding that reducing malnutrition by 5 percent could reduce under five child mortality by 30 percent. Although intuitively it seems to be clear that being malnourished increases the risk of child mortality, considerable doubts concerning the closeness of the relationship exist.

Assuming a close relationship between child mortality and undernutrition two glaring puzzles exist when the two regions of South Asia and Sub-Saharan Africa are compared. The first puzzle is the so called South Asian Enigma. The anthropometric outcomes are considerably better in Sub-

Saharan Africa than in South Asia. Almost half of the children in South Asia are malnourished. Compared to Sub-Saharan Africa the anthropometric shortfall is almost 70 percent higher in South Asia (WHO 2005), despite higher per capita calorie availability and better provision of health care, water and sanitation (Ramalingaswami et al 1996; Osmani 1997; Svedberg 2002). The second puzzle concerns the existing child mortality reversals between these two regions (Svedberg 1999; Svedberg 2000; Klasen 2003). In contrast to the severe anthropometric failure in South Asia, Sub-Saharan African countries suffer by far the worst from high rates of child mortality. In Sub-Saharan Africa 174 children out of 1000 die before reaching the age of five and in South Asia 97 (UNICEF 2004). Together, these two puzzles can then be defined as the South Asia Sub-Saharan Africa Enigma of anthropometric failure and mortality reversals.

There exist various possible explanation for the Enigma in the literature. First, clearly the level of income poverty is a major cause both for child mortality and undernutrition, but this cannot explain the regional differences because the average incidence of poverty is quite similar in the two regions. Second, it might be the way undernutrition is measured. For example, Klasen (1999) argues that the US-based reference standard for international comparison of undernutrition proposed by the WHO (1995) leads to an overestimation of undernutrition in South Asia. This overestimation could be due to different genetic potential in growth between the population in these two regions. The high level of undernutrition of children in South Asia might then appear because they are genetically shorter and/or lighter compared to the reference population and are therefore spuriously considered as malnourished. But even if this is the case, this could explain only a part of the huge differences in the anthropometric outcomes. Also the use of the new reference standard by the WHO that is based on child growth data from six different countries is in no way able to solve the enigma. Besides several authors have found evidence that there exist no real genetic differences between children's growth paths below the age of five in South Asia (see, for example Gopalan 1992; Eveleth and Tanner 1990; Svedberg 2000; Svedberg 2002) which suggests that these differences are caused by other factors, although a final conclusion concerning the influence of genetic factors on children's growth paths is not yet possible. A third aspect are clearly the different disease patterns between the two regions. The high incidence of HIV/AIDS and Malaria can potentially explain part of the enigma, but a further assessment of this effect is strongly constraint by data availability. Fourth, the primary health care provision and other public services are possible explanations which are less adequately provided in Sub-Saharan Africa (Svedberg 1999; Ramalingaswami et al 1996). Fifth, a further explanation is that the same determinants of child mortality and undernutrition may have different impacts in the two regions or that both phenomena are not as closely related as generally assumed.

Explaining the different relationships of child mortality and undernutrition between these two forms of deprivation within a country and also between countries and regions has considerably important policy implications because it helps to allow a much more detailed assessment of needed policy interventions and a better targeting to fight child mortality and undernutrition and to meet the MDGs. But approaches using aggregated macro data have not been able to explain it appropriately. Until now we find no attempts to explain the South Asia and Sub-Saharan Africa Enigma from a microeconomic perspective that have analyzed the socioeconomic determinants simultaneously for child mortality and undernutrition with the focus on their differences and similarities using micro data.

The aim of the paper is helping to explain the Enigma. To achieve this, we simultaneously try to find what socioeconomic determinants effect child mortality and undernutrition. In particular, we try to find out which determinants drive undernutrition as well as child mortality in a similar way and what factors have differing effects on both phenomena.

In contrast to most cross country studies made so far that investigate the determinants of child mortality and undernutrition, we introduce the methodology of multilevel modelling into our analysis that explicitly takes the hierarchical structure of the Demographic and Health Survey (DHS) data sets into account. This will also help to provide information about differences in the outcome variables due to differences in community characteristics especially about the provision of infrastructure service. We investigate the effects of individual, household and cluster socioeconomic characteristics on anthropometric shortfalls and child mortality and examine their respective influences and relationships on both phenomena and capture both within and between community effects in a single model. For the empirical analysis we use several nationally representative Demographic and Health Surveys (DHS) for a sample of six developing countries in South Asia and Sub-Saharan Africa. We find that the determinants of child mortality and undernutrition differ significantly from each other. Access to health infrastructure is more important for child mortality, whereas individual characteristics like wealth and educational and nutritional characteristics of mothers play a larger role for anthropometric shortfalls. Although very similar patterns in the determinants of each phenomenon are discernable, there are large differences in the magnitude of the coefficients. Besides regressions using a combined data set of all six countries show, that there are still significant differences between the two regions. Both region dummies as well as numerous interaction effects are significant. Therefore, given the underlying data and the proposed methodology, the South Asia - Sub-Saharan Africa Enigma can not be fully solved by different levels in access to health facilities, education, wealth, status of women alone.

The paper is structured as follows. After the given problem statement and an overview about the existing literature on measuring child mortality and child undernutrition and the differences in their outcomes in South Asia and Sub-Saharan Africa, section 2 explains the empirical method of multilevel models and specifies our model. Section 3 presents the data sources. In section 4, first descriptive statistics show the different patterns of child mortality and undernutrition within and between the analyzed countries. Second, we provide estimation results of the multilevel analysis. Third and finally, we simulate changes in the outcome variables for changes in selected covariates. Section 5 concludes and draws the policy implications from the results.

2 Empirical Approach

2.1 Multilevel Analysis

Many surveys in economics have a clustered or hierarchical data structure where a hierarchy consists of units grouped at different levels. For instance, individuals (level 1) are nested within households (level 2), households are nested within communities (level 3) and communities are even nested within states and countries. Standard regression models have problems dealing with the hierarchical data structure, even if we only include variables at level one (i.e., the child level), because they assume independent and normally distributed errors with a constant variance. Analyzing variables from different levels without taking into account the hierarchical data structure leads to misleading estimation results, because one faces the problem of heteroscedasticity. The individual observations in hierarchical data structure are not completely independent and the results of the analysis can be effected by this clustered structure of the underlying data. Put it differently, households in the same community are more homogenous than households in different communities. In particular, in the case of child undernutrition this means that the anthropometric outcomes in different communities might be independent from each other, but that outcomes within a community, especially when the children live in the same household. This leads to a violation of the assumption of independent errors which has consequences to the estimation results. The estimated coefficients are unbiased but not efficient because the standard errors are negatively biased which results in misleading significance effects. What is typically done in the empirical literature is to regress on independent variable at the lowest level on a set of explanatory variables available for any other levels by disaggregating all higher level variables to the individual level. This is done, for example, by assigning each individual in the same community the same value of the community variable. But this leads to the problem of inefficient estimation results mentioned before.[1]

In this analysis we want to study on the basis of clustered household surveys whether mortality rates and rates of undernutrition differ between several individual and household characteristics that vary from community to community. Furthermore, we are concerned with the understanding the factors associated with variations between countries and, within a country between communities. This means, we want to analyze the impact of community characteristics on the two outcome variables e.g., the access to health facilities and how much of the between community variation is explained by community explanatory variables. To conduct this study, in contrast to the use of standard regression models, a more adequate way to take the hierarchical data structure into account is the methodology of multilevel modelling. A multilevel model concerns the analysis of the relationship between variables that are measured at different hierarchical levels (Hox 2002).[2] The aim of a

[1] One can also think of aggregating the variables of the individual level to a higher level and do the analysis on the higher level. But this leads in many cases to a loss of the within-group information we are interested in.

[2] The first multilevel analysis in the social science was done by Aitkin et al (1981). They analyzed the impact of the teaching style on progress in reading capabilities of children in primary schools in Great Britain using traditional multiple regression techniques shown by Bennett (1976). When the data is analyzed only with the individual children as the units of the analysis without recognizing that they are groups within classes the results were statistically significant. When the grouping of children in classes is taken into account,

multi-level model is to take this data structure explicitly into account and to determine the direct effect of the individual and the group explanatory variables. Methodological work on analyzing multilevel models was done, for instance, by Bryk and Raudenbush (1992), Goldstein (1987, 1999) and more recently by Hox (2002), who gives an illustrative introduction in multilevel models with an application to educational data.

Using a multilevel model approach provides several advantages when analyzing clustered survey data because it allows the use of both individuals and groups of individuals simultaneously in the same model without violating the assumption of independent cases because the model includes the various dependencies between the variables. Multilevel models correct for the bias in the parameter estimates resulting from the clustered data structure because in a multilevel model each level is represented by it own sub-model which expresses the relationship among explanatory variables within that level. This possibility leads to several advantages using multilevel modelling. First, it provides statistically efficient estimates of the regression coefficients by providing correct standard errors, confidence intervals and significance tests (Goldstein 1999). Second, cross-level effects and cross-level interactions, i.e., the relationship of variables at different levels, can be analyzed. This means, measuring covariates at each level provides the possibility to analyze the extent to which differences in child mortality and undernutrition between communities are due to community factors like access to health facilities or due to factors at the individual level like gender. Third, estimates of the variances and covariances at each level of the model allows to decompose the total variance in the outcome variable into fractions for each level. In the so called variance component models the error term is divided into two parts, the group component and the individual component. This allows the assessment of the variation that is due to differences at the group level and

then the significant differences between teaching styles found before disappear.

due to differences at the individual level.[3]

2.2 The Basic Multilevel Model

In a multilevel model, the dependent variable is located at the lowest level, in our case the individual (child) level. Following Hox (2002) the basic multilevel model with two different levels can be described as follows. Suppose that we have $j = 1, ..., J$ level 2 units (i.e. communities) where there are $i = 1, ..., n_j$ level 1 units (i.e. children). Then we can speak of child i is nested within community j. In a multilevel model, the dependent variable is at the lowest level, in our case the individual (child) level. To analyze the outcome variable we can set up the regression equation as follows:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \tag{1}$$

with β_0 as the intercept and the slope β_1, defined as the expected change in the dependent variable with an increase in the individual variable X of one unit.[4] The difference to standard regression models in equation (1) is that there are two subscripts one referring to the individual i and one to the community level j. The clustered data structure and the within and between community variations is now taken into account by assuming that each community has a different intercept β_{0j} and a different slope β_{1j}. Then the explanatory variables at the second level Z can be introduced in the model. For this the coefficients β_{0j} and β_{1j} are themselves given in a regression model as dependent variables via two regression equations with the level two variables as the independent explanatory variables:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \tag{2}$$

[3]For instance, Pebley et al (1996) investigates the receipt of vaccinations of children in Guatemala with variables at the individual, at the household and at the community level. When controlling for the observed variables, they found that the variance due to households is five times higher than due to communities.

[4]We assume that the errors e_{ij} have a mean of zero so that $E(e_{ij}) = 0$ and a variance $var(e_{ij}) = \sigma_e^2$ so that $e_{ij} \sim N(0, \sigma_e^2)$.

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j}. \tag{3}$$

Equation (2) and (3) explain the variations between communities because the intercept β_{0j} and the slope β_{1j} depend on the community variables in community j. For example, equation (2) predicts the average anthropometric outcome of the child by the level 2 variable Z in community j. Equation (3) states that the slope β_{1j} between the anthropometric outcome (Y) and level 1 variable (X), i.e. gender, depends on the level 2 variable (Z), i.e. access to health. The error terms u_{0j} and u_{1j} are level 2 residuals.[5]

The combined model can no be written by one single complex regression equation by substituting (2) and (3) into (1):

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} X_{ij} Z_j + (u_{1j} X_{ij} + u_{0j} + e_{ij}). \tag{4}$$

In a more general form, assuming that we have P explanatory variables X at the lowest level, denoted by the subscript $p(p = 1...P)$ and Q explanatory variables Z at the highest level, indicated by the subscript $q(q = 1...Q)$ equation (4) becomes to:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + \gamma_{pq} X_{pij} Z_{qj} + (u_{pj} X_{pij} + u_{0j} + e_{ij}). \tag{5}$$

In equation (5) the first part can be defined as the deterministic part referring to the fixed coefficients, which means, that coefficients do not vary across level. The part of equation (5) expressed in brackets can be defined as the stochastic part, containing the random error terms. The term $X_{ij}Z_j$ is an interaction term analyzing the cross-level interaction.[6]

[5] The residuals u_0j and u_{1j} are also assumed to have mean of zero so that $E(u_{oj}) = E(u_{1j}) = 0$. It is also assumed that the variance is defined as $var(u_{oj}) = \sigma^2_{u0}$, $var(u_{1j}) = \sigma^2_{u1}$, and the covariance as $cov(u_{oj}, u_{1j}) = \sigma_{u01}$. A positive value of the covariance between β_0 and β_1 indicates that communities with high means tend also to have positive slopes. In addition, it also assumed that level 1 residuals are not correlated with the level 2 residuals so that $cov(u_{oj}, e_{ij}) = cov(u_{1j}, e_{ij}) = 0$.

[6] As OLS estimations techniques are inappropriate to deal with the within level two dependencies, the multilevel analysis is mostly based on an iterative maximum likelihood estimation (Mason et al 1983, Goldstein 1987, Bryk and Raudenbsuh 1992). An advantage of the maximum likelihood method is that it provides estimates that are asymptotically efficient and consistent (for a detailed description of maximum likelihood estimation technique see, e.g., Eliason 1993).

The stochastic part in equation (5) demonstrates again the problem of dependent errors. In contrast to standard ordinary least squares (OLS) regression the error term in (5) contains one individuals component e_{ij} and a group or community component $u_{0j} + u_{1j}X_{ij}$. The individual error component e_{ij} is independent across all individuals. In contrast, the community level errors u_{0j} and u_{1j} are independent between communities but dependent within each community because the components are common for every child i in community j. These dependencies lead to unequal variances of the error terms which results into heteroscedasticity, because $u_{0j} + u_{1j}X_{ij}$ depend on u_{0j} and u_{1j} which vary across communities and on X_{ij} which vary across children.

2.3 Model Specification

In our multilevel analysis we set up a two-level model. The level one includes both individual and household variables and the second level is the cluster level. We do not separate between the individual (child) level and the household level, because there are no real differences between individual and the household information, because there are only a very few households with more than three young children in the data.[7]

The empirical analysis proceeds in 6 basic steps. First, we run several regression model types to get a benchmark for our two outcome variables and to explain the differences between the multilevel approach and standard regression models. For child mortality we run both a proportional hazard model (Cox and Oakes 1984) and logit regression. For stunting we also run a logit regression on a dummy whether the child is stunted and a OLS regression on the stunting z-scores. In the second step, to built up the multilevel model, we start by including all explanatory variables of level 1 into the model which means that the variance component of the slopes are

[7] When setting up a multilevel model, Mass and Hox (2004) suggest a sample size of the second level of more than 50.

fixed to zero.[8] This model serves us as a benchmark for the two variance components. Third, we set up the full model by adding the explanatory variables of the community level. Comparing this model with the model in step three allows us to investigate whether and to what extent the between community variation in child mortality and child undernutrition is explained by community characteristics.

For a meaningful interpretation of the intercept we center each explanatory variable around the grand mean by subtracting the grand mean from each variable.[9] Thus, equation (5) is changed to:

$$Y_{ij} = \gamma_{00} + \gamma_{p0}(X_{pij} - \bar{X}_p) + \gamma_{0q}(Z_{qj} - \bar{Z}_q) + \gamma_{pq}(X_{pij} - \bar{X}_p)(Z_{qj} - \bar{Z}_q) + [u_{pj}(X_{pij} - \bar{X}_p) + u_{0j} + e_{ij}]. \quad (6)$$

So far we have described the multilevel model assuming continuously distributed dependent variables (i.e., income or stunting z-scores). However, when analyzing child mortality the dependent variable is a proportion or a dummy variable. The same holds for the proportion of stunted children. Thus the two-level model described in equation (6) for binary data, where $Y_{ij} = \pi_{ij}$ can be written as follows:

$$logit(\pi_{ij}) = \gamma_{00} + \gamma_{p0}(X_{pij} - \bar{X}_p) + \gamma_{0q}(Z_{qj} - \bar{Z}_q) + \gamma_{pq}(X_{pij} - \bar{X}_p)(Z_{qj} - \bar{Z}_q) + [u_{pj}(X_{pij} - \bar{X}_p) + u_{0j}]. \quad (7)$$

Equation (7) is quite similar to equation (6), but the outcome variable is now a proportion.[10]

[8] In particular, we assume that $u_{pj} = 0$.

[9] The reason of centering the explanatory variables is the interpretation of the intercept β_0. As it is defined as the expected value of the outcome variable when all explanatory variables have a value of zero, we face the problem that this would be misleading for some dummy variables because they are coded as 1 and 0. If we center the variables around their grand mean, the intercept becomes the expected value of the outcome variable, when all variables have their mean value (e.g., it becomes the mean z-score).

[10] In particular, it is assumed that π_{ij} has a binomial error distribution with expected value μ and sample size n_{ij}, so that $\pi \sim Bin(n_{ij}, \mu)$ and the variance $var(\mu_{ij}) = \sigma^2 = (\pi_{ij}(1 - \pi_{ij}))/n_{ij}$. Note that equation (7) contains no error term e_{ij}. Because the errors

After the multilevel analysis for each country in the sample to identify differences in the effects of the explanatory variables on child mortality and undernutrition between countries in South Asia and Sub-Saharan Africa, in step five we merge all country data sets to on global data set and run again the multilevel regression asking for specific country and region fixed effects. This done in two ways. First, we include country dummies into the regression to identify country differences in the covariates. Second, a Sub-Saharan Africa dummy is included to capture regional differences. In addition, the dummy is also interact with all explanatory variables at each level. Finally, in step six the previous analysis is extended by constructing a simulation of several scenarios for child mortality and undernutrition. Here, we compare changes in the outcome variables for potential changes in specific covariates.

3 Data

To obtain possible explanations about the regional differences in child mortality and undernutrition between South Asia and Sub-Saharan Africa, we analyze a sample of five countries from these regions. We use national representative demographic and health surveys (DHS) surveys that provide information on anthropometric outcomes of the children, information about access to the health system and other information about the socioeconomic status of children below the age of five and the mothers (aged between 15 and 49). The DHS data do also contain information on cluster characteristics, especially on infrastructure. This information is included in the service availability recodes that are in our case available for the South Asian countries Bangladesh (2000) and India (1999) and in Sub-Saharan Africa for Mali (2001), Uganda (1995) and Zimbabwe (1994). With this country data sets,

are binomially distributed, the residual error variance is a function of the population proportion $\pi_{ij} : \sigma^2 = (\pi_{ij}(1 - \pi_{ij}))/n_{ij}$ and it is therefore not necessary to estimate it separately (Hox 2002).

the sample contains more than 53.000 children in South Asia and more than 29.000 children in Sub-Saharan Africa that enter the analysis.

As dependent variables to study child mortality and undernutrition we use two dummy variables. For child mortality the dummy whether the child died in the first year of life.[11] To measure child undernutrition, the DHS data sets provide information on several anthropometric outcomes of children, in particular the z-scores for weight for age, weight for height and height for age.[12] We use a dummy whether the child is stunted that is if the stunting z-score (height for age) is below -2 standard deviation from the median of the reference population (WHO 1995).[13]

In the empirical model we include a set of several individual and household characteristics as well as cluster characteristics that might have an effect on the two outcome variables. For the individual characteristics, besides the household size and the number of children in the household, we include the age and sex of the child into the regression equation. The rate of undernutrition is supposed to decrease with increasing age of the child and with the sex variable we control for sex differentials in mortality and undernutrition in our countries as it is often be found in the empirical literature concerning child mortality and undernutrition (for example, see Marcoux 2002 and Klasen 1996). Other major determinants especially on child mortality are the preceding birth interval of the mother and the question whether and when they child was breast fed. Breastfeeding in the first month of life plays an important role for the development of the child because the breastmilk

[11]To capture the whole birth history of the children, we do not consider child mortality of children below the age of five because this throws out to many observations. We do not explicitly separate between neonatal deaths (child died in the first month) and post-neonatal death (child died between the first month and the first year of life proposed, for example by Adebayo et al (2004) because this did not change the results.

[12]For example, the stunting z-scores are the outcome of the ratio of height over age minus the median of the reference population and the standard deviation of the reference population (see, e.g.,Klasen 1999; Smith and Haddad 2000).

[13]We also consider the case of extreme stunted children where the z-score is below -3 standard deviation of the height for age norm.

meets most of the child's needs and makes the child more resistent against diseases (Ramalingaswami et al 1996). Concerning the mother, the educational level of the mother enters the regression equation. The argument here is twofold. First, more educated women might be better able to process information and acquire skills to take care of the children, for example in the case of illness, and second, better educated women are more able to earn money. In addition, the nutritional status of the mother is included, which is supposed to strongly effect the nutritional status of the child.[14]

As we do not have information on income or expenditure in the DHS surveys we consider an asset-based approach in defining well-being (Sahn and Stifel 2001). For this we use a factor analysis on several household assets proposed by Filmer and Pritchett (2001) to derive an index that indicates the material status of a household. This approach has proved to be a good approximation of wealth. In particular, as the weights for the asset index we include dummies whether the following assets exist or not: Radio, TV, Refrigerator, Bike, Motorized transport, Low floor material, Toilet, Drinking water. Then we introduce another index into the analysis that includes information about the access to health facilities of the household. Again this is based on a factor analysis asking whether the mother has received a tetanus vaccination before birth, whether the mother has received prenatal care, whether the child was born at home without assistance of a doctor or a nurse and also the average number of vaccinations per child within a household. We assume that the access to health facilities is an crucial determinant both for child mortality and undernutrition. This index provides an additional advantage because it captures both the potential access opportunities to the health system and is also outcome really which means, that the child or the mother have really benefited from the service.

Besides the individual and household characteristics we include cluster

[14]The recommend method to measure the nutritional status of adults is the body mass index.

variables.[15] In this context, the multilevel model distinguishes two different kinds of variables, namely contextual variables and global variables. Contextual variables at higher levels are variables that are simply the aggregates of the covariates at the individual level for each cluster. For example, we include the percentage of women with secondary education per cluster and the percentage of children that had recently suffered from fever per cluster. The global variables are part of the service availability recode and are not drawn from information of the individual level. In our case these global variables provide information about the infrastructure in the cluster. We include the distance to the next health facility which might be important for the access to heath services and a public infrastructure index that is based on the availability of general facilities like a bank, a cinema, a post office etc. The weights are again determined by a principal component analysis.

4 Results

In this section we first show some descriptive statistics that are relevant for the following multivariate regressions of child mortality and stunting. The section is concluded by a short discussion of simulations describing the economic significance of different determinants of both phenomena.

4.1 Descriptive Statistics

As seen in Table 1 the South Asia Sub-Saharan Africa Enigma of anthropometric failure and mortality reversals is clearly discernable in our six data sets. Higher undernutrition rates in both South Asian countries coincide with lower infant mortality rates than in the three Sub-Saharan African countries. This result is independent of the measure for undernutrition (i.e. stunting, wasting, underweight or the Composite Index of (Severe) Anthropometric Failure (CIAF/CISAF) that indicates undernutrition by any of the

[15] In the case of India the service availability recode contains information by districts instead of cluster.

preceding measures). This picture is not changed by the use of the new multicountry growth reference standard that was published by the WHO in 2006. Prevalence rates using this new reference standard are shown in parentheses.

[please insert Table 1 here]

While the number of possible determinants of child mortality and undernutrition is extremely large the following section is focused on determinants that are known to have a significant influence on child mortality and undernutrition. Covariates that had no significant influence in our numerous model specifications were dropped.[16]

[please insert Table 2 here]

As shown in Table 2 there are large differences in the covariates between countries. But these differences seldom form clear regional patterns. One clear exception is the status of mothers, which is a lot worse in South Asia than in Sub-Saharan Africa. On the one hand the percentage of undernourished mothers is three to five times higher in South Asia. On the other hand the age at marriage and at first birth are lower and the number of unwanted children larger especially in Bangladesh, which are strong indicators of stronger gender discrimination.

As mentioned before the lack of income data necessitate the use of a wealth index as a proxy for incomes and consumption. To avoid using arbitrary weights we use a principal components analysis, which means that the weights are equivalent to a measure of the degree of correlation between each factor and a hidden component (i.e. in our case wealth).

As seen in Table A3 the weights for the factors have the assumed sign, giving positive values to durable goods like TV and radio and negative values to the lack of a toilet facility or the use of surface drinking water.

[16] One example for such a variable without significant influence on both phenomena is the sex of a child.

Also when we look at the weights of our Health Facility Index it can easily be seen that the principal component analysis determines weights with the "right" signs. Therefore positive weights are generated for the dummies for a tetanus vaccination of the mother before birth and for prenatal care as well as for the mean number of vaccinations per child in household. A negative value is generated for the dummy whether a child was born at home without the assistance of a doctor or a nurse.

Both factors wealth and access to health facilities that are proxied by our indices are strongly correlated with child mortality and undernutrition. As seen from Tables A1 - A3 both phenomena are a lot more prevalent in the lower quintiles of both indices. A particularly strong connection is observable between access to health facilities and child mortality.

4.2 Regression Results

As mentioned before we use a multilevel model approach to examine the influence of individual, household and cluster socioeconomic characteristics on child mortality (Tab. 3) and undernutrition (Tab. 4). The use of a multilevel approach instead of a standard regression models insures that we avoid misleading significance effects due to violations of the assumption of independent errors with a constant variance. This effect is confirmed in our regression results, in which the multilevel regressions display lower levels of significance compared to the OLS regression, the proportional hazard regression and the logit regression with the same model specification. Especially in the case of community characteristics a strong reduction in significance levels is observable.

Tables 3 and 4 show that child mortality and undernutrition have very similar determinants across countries. Although there are considerable differences in the magnitude of the coefficients both significance and direction of the influence conform in the majority of cases.

[please insert Table 3 here]

Age has in all cases a significant non-linear negative influence on child mortality meaning that the number of child deaths decreases non-linearly with age. At the same time age influences undernutrition positively in a well known non-linear way as shown in Fig. A1. Very similar effects across countries are also found when looking at other individual characteristics like immediate breastfeeding, the birth interval to the preceding birth and the dummy variable for being first born or not. As expected a positive feeding practice like the immediate initiation of breastfeeding after birth has a negative and in most cases significant effect on infant mortality. This complies with the general knowledge on the importance of the colostrum that contains a large number of antibodies and basically works as a first immunization or vaccination. Similarly birth spacing results in a significant reduction in child mortality. On the other hand being the first born increases the risk of dying within the first year in all countries, which could be due to a lack of experience of mothers in nurturing a child or in the recognition of illnesses.

Analogous to the individual characteristics we find very consistent patterns of the determinants at the household level. A higher fertility, measured by the total number of children born by a mother up to the date of the interview, increases the mortality risk significantly. Contrary to that the household size has a mortality reducing effect, possibly reflecting the better capability of larger households to cope with shocks and/or a larger stock of knowledge on raising children within a household. A positive influence is also exerted by the status of women which is proxied by the age of mothers' at marriage, but the effect is only significant in two of the six countries. Quite surprisingly other important characteristics of the mothers exhibit a much lower influence on child mortality than originally expected. Taking account of the other determinants, a mother's level of education, measured as the amount of schooling in years, has no significant mortality reducing effect. Therefore a mother's education seems to have no secular influence on child mortality but influences it only via other determinants like better feeding practices and lower fertility that are separately considered in our model specification. At the same a bad nutritional status of the mother has a significant positive effect on the mortality risk of a child in only one country.

Even more surprising is the low separate influence of wealth, measured by our asset index, on reducing the probability of a child to die within the first year. It is even the case that in five countries the regression coefficients of the asset index are positive and in three cases these coefficients are significant.

By far the largest and most significant effect on child mortality is exerted by the access to health facilities, which is measured by our health facility index. This index includes information on whether the mother received a prenatal care as well as a tetanus injection before birth, whether the child was born at home without the assistance of a doctor or a nurse and on the mean number of vaccinations per child in a household. This effect is not limited to a high level of statistical significance, but our simulations show that the level of economic significance is also very high.

Opposite to that, we find no strong determinants of child mortality at the community level when using the multilevel approach. Neither the distance of a cluster to the next health facility, the percentage of mothers with secondary education, an index of public infrastructure, nor the percentage of children with fever in a cluster, nor any other variable tested at the community level have significant effects on child mortality. This is the case although the variation of the intercept of the community level σ_{u0}^2 is significant and therefore shows that information on this level plays a role in explaining child mortality. Nevertheless by including those four variables we can explain a significant part of the variation at the community level. Additionally those explanatory variables at the community level have the expected significant effects in most cases when they are included in a regression model without consideration of variables at the individual and household level. As soon as these variables are included the community characteristics lose their significance.

[please insert Table 4 here]

Comparing the determinants of child mortality with those of undernutrition we find that differences are larger than expected. The opposing influence of age on both phenomena was already mentioned. While immediate breastfeeding helps to reduce the risk of undernutrition, its effect is much lower and less significant than in the case of infant mortality. Being the first born is also much less detrimental to the nutritional status of a child. Quite the contrary in five of the six countries the coefficient of the variable is negative and in three cases also significant, showing that the lack of experience in raising children is not a significant factor influencing the nutritional status of a child in a negative way.

At the household level the differences in the determinants of child mortality and undernutrition are especially large. The only factor having a similar influence is the access to health facilities. This clearly reduces the incidence of undernutrition in a significant way, although the magnitude and significance of the effect is of a much lower scale. In contrast to that the wealth of a household helps to significantly reduce the probability of being undernourished. In addition to that individual characteristics of the mother like her level of education, a higher age at marriage and a good nutritional situation all significantly improve the nutritional status of a child. The household size has contrary to infant mortality no reducing effect on undernutrition, possibly reflecting an offsetting of the better capability of coping with shocks by the larger amount of competitors for the limited household resources.

At the community level there is only one variable that seems to be of some significance, which is the education of mothers. As in the case of infant mortality the variation of the intercept of the community level σ_{u0}^2 is significant and the inclusion of the four community characteristics improves the goodness of fit significantly.

[please insert Table 5 here]

The additional regressions that were implemented using a combined data set of all children in the six countries confirm the results of the country regressions and show that there are still significant differences between the two regions even when we control for the large set of explanatory variables. The first row in Tab. 5 shows that child mortality is significantly larger in Sub-Saharan Africa than in South Asia and the second third row shows that it's the other way round when we look at stunting. Besides the significant region dummy the inclusion of region interaction effects shows that the coefficients for almost all variables differ significantly between regions. For example the positive influence of the access to health facilities in reducing child mortality and undernutrition is significantly lower in Sub-Saharan Africa than in South Asia, pointing to a possibly lower quality of health facilities in Sub-Saharan Africa. The interaction of the variable of age at marriage with the SSA-Dummy shows that improvements in the status of women will have a potentially larger effect in South Asia.

4.3 Simulations

Adding to our multilevel regressions we simulated a large set of equalizations and assimilations in different covariates in the two regions. Non of these simulations had the potential to fully explain the South Asia Sub-Saharan Africa enigma. This is not very surprising since the averages in the parameter values do not differ a lot between regions (although they differ quite a lot between countries).

But using these simulations we were able to test the economic significance of the different explanatory variables, meaning we were able to see what effects certain improvements in the different determinants have on both phenomena. One clear result was that changes in explanatory variables would result in very different changes in the two deprivations. The strongest influence on child mortality is exerted by the access to health facilities proxied by our health facility index. Although the influence on stunting was also very significant, the magnitude was by far not as large. Therefore improvements in health facilities will help a lot more in reducing child mortality than improving the nutritional status of children. At the same time we confirmed the preceding results that increases in wealth/income will result in significant reductions of undernutrition. Even stronger improvements in the incidence of undernutrition could be generated by increases in the level of education of mothers, that has no significant positive effect on changes in mortality rates on its own. Although immediate breastfeeding is statistically significant for child mortality and undernutrition feeding all children immediately after birth would only result in a economically significant reduction in child mortality.

5 Conclusion

In the preceding analysis we investigated the effects of individual, household and cluster socioeconomic characteristics on child mortality and undernutrition using a multilevel model approach. We find strong evidence in support of the existence of the South Asia Sub-Saharan Africa enigma using micro data. While generally having very similar patterns across countries, we find that the determinants of child mortality and undernutrition differ significantly from each other. Access to health infrastructure is more important for child mortality, whereas the individual characteristics like wealth and educational and nutritional characteristics of mothers play a larger role for anthropometric shortfalls. Although very similar patterns in the determinants of each phenomenon are discernable, there are large differences in the magnitude of the coefficients. Besides regressions using a combined data set of all six countries show that there are still significant differences between the two regions although taking account of a large set of covariates. While the average parameter values are quite similar in the two regions, it can be shown by including interaction effects between regions and the different explanatory variables in the regressions that the size of the coefficients varies significantly between regions.

One hypothetical explanation for the regional differences remains in the quality of the data. There might be biases and errors especially in the African data sets. But these biases cannot account for the differences in the determinants of both phenomena, since the same data sets and explanatory variables are used for the explanation of child mortality and undernutrition in all countries. Another possible explanation for the enigma might lie at least in part in the different occurrence of diseases like HIV/AIDS and Malaria. Further studies will therefore try to estimate the impact of HIV/AIDS on infant mortality rates. Further aspects of future research could try to capture differences in the quality of health facilities and in the composition of foods.

As our study has also shown there are determinants at the cluster level that have a significant influence on child mortality as well as undernutrition. Unfortunately the available variables were not able to capture this information. Therefore additional research could try to detect variables at this level.

Finally part of the explanation of the South Asia Sub-Saharan Africa enigma could be the insight that child mortality and undernutrition are not as closely correlated as generally assumed. Our study finds considerable evidence for large differences in the determinants of both phenomena. These differences make it highly unlikely that child mortality and undernutrition are as closely correlated as found by the studies of Pelletier et al. (1995, 2002) and cited by numerous other publications.

Therefore it could be more difficult to achieve both Millennium Development Goals concerning child mortality and undernutrition. The implementation of a set of policies that will help to reduce both phenomena in the same way seems to be illusionary. On the contrary it seems to be necessary to find an appropriate set of policies for each of the two deprivations. Child mortality reductions can mainly be achieved by improvements in public health infrastructure. Although this will also help to reduce the incidence of undernutrition the effect won´t be as large. Contrary to that improvements in gender related aspects like the education of mothers and the status of women will contribute significantly more to declining numbers of undernourished children. Major improvements in undernutrition rates can also be achieved by increases in personal wealth.

References

Adebayo, S., L Fahrmeir and S. Klasen (2004), Analyzing Infant Mortality with Geoadditive Categorical Regression Models: A Case Study for Nigeria, *Economics and Human Biology* 2: 229-244.

Atkin, M, D. Anderson and J. Hinde (1981), Statistical modelling of data on teaching styles (with discussion), *Journal of the Royal Statistical Society* 144: 148-161.

Bennett, N. (1976), *Teaching Style and Pupil Progress*, London, Open Books.

Bryk, A.S. and S.W. Raudenbush (1992), *Hierarchical Linear Models: Applications and Data Analysis*, Newbury Park, CA, Sage.

Cox, D.R. and Oakes, D. (1984), *Analysis of Survival Data*, New York.

Eveleth, P.E. and J.M. Tanner (1990), *Worldwide Variations in Human Growth*, Cambridge, Cambridge University Press.

Filmer, D. and L.H. Pritchett (2001), Estimating Wealth Effects without Expenditure Data - or Tears: An Application to Educational Enrollments in States of India, *Demography* 38 (1): 115-132.

Gopalan, C. (1992), Undernutrition: Measurement and Poverty, in *Undernutrition and Poverty*, ed. Osmani, S.R., Oxford.

Goldstein, H. (1999), *Multilevel Statistical Models*, London, Arnold.

Goldstein, H. (1987), *Multilevel Models in Educational and Social Research*, London, Griffin.

Eliason, S.R. (1993), *Maximum Likelihood Estimation*, Newbury Park, CA, Sage.

Gorad, S. (2003), What is Multi-Level for?, *British Journal of Educational Studies* 51 (1): 46-63.

Hox, J. (2002), *Multilevel Analysis - Techniques and Applications*, Mahwah, NJ, Erlbaum.

Klasen, S. (2003), Malnourished and Surviving in South Asia, better Nourished and Dying Young in Africa: What can Explain this Puzzle? in: FAO (eds.) Measurement and Assessment of Food Deprivation and Undernutrition, pp. 283-286, Rome: FAO.

Klasen, S. (1996), Nutrition, Health, and Mortality in Sub-Saharan Africa: Is there a Gender Bias?, *Journal of Development Studies* 32: 913-933.

Klasen, S. and A. Moradi (2000), The nutritional Status of Elites in India, Kenya and Zambia: An appropriate guide for developing reference standards for undernutrition? Sonderforschungsbereich 386: Analyse diskreter Strukturen Discussion Paper No. 217.

Maas, C.J.M. and J. Hox (2004), Robustness issues in multilevel regression analysis, *Statistica Neerlandica* 58 (2): 127-137.

Marcoux, A. (2002), Sex Differentials in Undernutrition: A look at Survey Evidence, *Population and Development Review* 28 (2): 275-284.

Mason, W.M., G.M. Wong, and B. Entwistle (1983), Contextual analysis through the multilevel linear model, *Sociological Methodology* 13: 72-103.

McCullagh, P. and Nelder, J. (1989), *Generalized Linear Models (2nd edition)*, London, Chapman and Hall.

Osmani, S.R. (1997), Poverty and Nutrition in South Asia, in:*Nutrition and Poverty: papers from the ACC/SCN 24th Session Symposium*, Kathmandu, March 1997, United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition, New York.

Pebley, A.R., N. Goldman, and G. Rodriguez (1996), Prenatal and delivery care and Childhood immunization in Guatemala: do family and community matter? *Demography* 33: 231-247.

Pelletier, D.L. and E.A. Frongillo (2002), Changes in Child Survival are strongly Associated with Changes in Malnutrition in Developing Countries, *The Journal of Nutrition*, 133: 107-119.

Pelletier, D.L., E.A. Frongillo, Jr, D.G. Schroeder and J.P. Habicht (1995), The effects of mulnutrition on child mortality in developing countries, *Bulleting of the World Health Organization* 73 (4): 443-448.

Pritchett, L. and L.H. Summers (1996), Wealthier is healthier, *Journal of Human Ressources* 31 (4): 841-868.

Ramalingaswami, V., U. Johsnon, and J. Rohde (1996), The Asian enigma, in: *Progress of Nations.* New York: United Nations Children´s Fund.

Sahn, D.E. and D. Stiefel (2003), Exploring Alternative Measures of Welfare in the Absense of Expenditure Data, *Review of Income and Wealth* 49 (4): 463-489.

Smith, L.C. and L. Haddad (2000), Explaining Child Malnutrition in Developing Countries, International Food Policy Research Institute, Washington.

Ssewanyana, S. and S. Younger (2004), Infant Mortality in Uganda: Determinants, Trends, and the Millennium Development Goals, Development Policy Research Unit (DPRU), University of Cape Town, South Africa.

Subbaro, K. and L. Rany (1995), Social gains from female education: A cross-national study, *Economic development and cultural change* 44 (1): 105-128.

Svedberg, P. (2002), Hunger in India - Facts and Challenges. Institute for International Economic Studies Seminar Paper No. 699, Stockholm.

Svedberg, P. (2000), *Poverty and Undernutrition: Theory, Measurement and Policy*, Oxford University Press, Oxford.

UN (2005), *The Millennium Development Goals Report 2005*, United Nations, New York.

UNICEF (2004), Progress for Children - A Child Survival Report Card Vol 1, UNICEF, New York.

WHO (2005), *Health and the Millennium Development Goals*, World Health Organization, Geneva.

WHO (2002), *World Health Report 2001*, World Health Organization, Geneva.

WHO (1995), Physical Status: The Use and Interpretation of Anthropometry, WHO Technical Report Series No. 854, World Health Organization, Geneva.

Tables

Table 1
Infant mortality and anthropometric indicators
(percentage)

	Bangladesh 2000	India 1999	Mali 2001	Uganda 1995	Zimbabwe 1994
Infant mortality					
Infant mortality	7.99	7.85	14.91	9.90	7.52
Undernutrition*					
Stunting	44.12	43.17	37.61	35.19	22.24
	(50.82)	(50.75)	(43.06)	(42.37)	(30.09)
Wasting	10.50	14.99	10.97	5.11	5.56
	(13.47)	(20.09)	(13.83)	(7.34)	(7.30)
Underweight	47.32	43.77	34.14	23.38	16.30
	(41.48)	(40.62)	(30.62)	(19.81)	(12.54)
Severe stunting	18.12	21.39	19.01	13.26	6.18
	(24.39)	(30.31)	(24.44)	(19.20)	(10.28)
Severe wasting	1.05	2.84	1.72	0.85	0.86
	(3.36)	(8.49)	(5.48)	(3.02)	(2.99)
Severe underweight	13.12	15.92	11.66	6.12	3.26
	(14.41)	(18.21)	(12.97)	(7.19)	(3.39)
CIAF**	56.63	57.21	47.86	41.21	29.82
	(61.10)	(63.58)	(52.05)	(47.01)	(36.09)
CISAF**	22.16	27.33	22.85	15.39	8.05
	(29.64)	(38.37)	(29.42)	(22.37)	(13.26)

Note: *Children are considered as wasted, stunted or underweight if the respective z-scores are below -2 standard deviation from the median of the reference category. If the z-scores are below -3, children are considered as severely undernourished. The numbers in brackets refer to the new reference standard for child anthropometric failure that was published by WHO in 2006. **CIAF and CISAF refer to the Composite Index of (Severe) Anthropometric Failure that indicates whether a child is (severely) undernourished by either stunting, wasting or underweight

Table 2
Summary statistics for individual, household and community characteristics

	Bangladesh 2000	India 1999	Mali 2001	Uganda 1995	Zimbabwe 1994
Number of Children	6.944	46.569	14.328	5.799	2.438
Individual characteristics					
Age (month)					
Mean	28.79	17.14	28.56	22.63	17.50
Breastfeeding*					
Child was breast fed	35.05%	25.81%	42.40%	59.13%	53.08%
Household characteristics					
Household size					
Mean	6.79	7.41	7.35	6.59	6.89
Total no. children	3.11	2.90	4.67	4.16	3.54
Household head					
Female	5.36%	6.53%	8.70%	20.28%	32.77%
Household has					
TV	17.98%	38.37%	16.07%	5.87%	12.29%
Radio	31.52%	41.73%	73.06%	49.19%	43.34%
Flush Toiled	10.93%	25.32%	7.19%	3.21%	22.38%
Piped Drinking Water	6.37%	40.04%	27.06%	12.58%	31.27%
Mother's education (years)					
Mean	3.19	3.90	0.91	4.13	6.38
(Standard deviation)	(3.78)	(4.71)	(2.45)	(3.52)	(3.68)
No education	45.36%	50.14%	83.68%	26.38%	12.92%
Primary education	28.96%	16.21%	11.33%	57.76%	51.27%
Secondary education	21.18%	24.45%	4.62%	15.67%	34.58%
Age at first marriage					
Mean	15.10	17.53	16.32	17.14	18.32
Age at first birth					
Mean	17.67	19.35	18.21	18.11	18.96
Child not wanted	16.30%	9.89%	4.72%	9.98%	9.31%
BMI of mother					
BMI<18.5	41.63%	34.85%	8.33%	7.77%	5.28%
Community characteristics					
Number of vaccinations					
Mean	5.44	4.65	3.83	5.03	5.82
Birth assistance					
Assistance at birth**	14.31%	44.39%	22.43%	44.16%	67.80%
Prenatal care	22.51%	62.88%	20.88%	90.46%	93.55%
Tetanus vaccination	62.37%	75.79%	31.37%	81.25%	82.53%
Born home w/o assist.	85.24%	55.08%	60.85%	22.77%	31.17%
Distance to health facility***					
Mean	47.67****	10.05	7.87	8.69	16.19
Children with fever recently					
Mean	37.81%	30.27%	31.08%	48.17%	37.73%

Notes: *Child was breastfed immediately after birth. **By doctor or nurse. ***Distance to hospital and clinic in kilometers. ****Time in minutes to next health facility is used instead of distance.

Table 3
Multilevel-Regression of infant mortality
(full model)

	Bangladesh 2000	India 1999	Mali 2001	Uganda 1995	Zimbabwe 1994
Fixed Part					
Constant	-3.658**	-2.764**	-2.345**	-1.895**	-2.234**
	(0.123)	(0.043)	(0.063)	(0.103)	(0.192)
Age of mother	-0.321**	-0.309**	-0.220**	-0.219**	-0.028
	(0.066)	(0.033)	(0.035)	(0.073)	(0.155)
(Age of mother)^2	0.449**	0.364**	0.260**	0.202*	-0.084
	(0.111)	(0.054)	(0.053)	(0.117)	(0.248)
Sex of child	-0.250*	-0.051	-0.130*	-0.087	-0.254
	(0.108)	(0.046)	(0.059)	(0.111)	(0.229)
Breastfeeding	-0.419**	-0.505**	-0.242**	-0.130	-0.336
	(0.124)	(0.067)	(0.064)	(0.116)	(0.231)
First born	0.472*	0.535**	0.194	0.385*	0.530
	(0.209)	(0.090)	(0.122)	(0.215)	(0.440)
Preceding birth interval	0.000	0.005**	-0.006**	-0.002	0.018**
	(0.004)	(0.002)	(0.002)	(0.004)	(0.006)
Household size	0.014	0.054**	0.006*	0.041*	-0.022
	(0.020)	(0.006)	(0.003)	(0.020)	(0.038)
Total no. children	0.117*	0.219**	0.154**	0.179**	0.094
	(0.051)	(0.024)	(0.023)	(0.046)	(0.112)
Asset index (global)	0.252*	0.005	0.023	-0.022	0.002
	(0.107)	(0.027)	(0.046)	(0.109)	(0.151)
Mother's education (years)	-0.010	0.012	-0.036*	-0.014	0.015
	(0.021)	(0.008)	(0.019)	(0.021)	(0.043)
Age at marriage	0.093**	0.081**	0.027*	0.070**	0.019
	(0.025)	(0.013)	(0.012)	(0.023)	(0.048)
Mother's BMI<18.5	0.000	-0.148**	-0.230*	0.214	-1.234
	(0.111)	(0.050)	(0.111)	(0.197)	(0.772)
Health facility index (global)	-1.588**	-0.864**	-0.535**	-1.112**	-1.230**
	(0.104)	(0.032)	(0.053)	(0.090)	(0.154)
Community characterisitcs					
Distance to health facility***	-0.003*	-0.009*	-0.001	0.003	0.007
	(0.002)	(0.004)	(0.003)	(0.005)	(0.007)
Percent children with fever	0.729*	-0.137	0.349	0.765*	0.580
	(0.377)	(0.371)	(0.231)	(0.326)	(0.611)
Percent secondary education	-0.213	0.446	0.542	0.229	-0.430
	(0.454)	(0.298)	(0.505)	(0.505)	(0.694)
Public infrastruct. index	0.036	-0.039	0.039	0.205*	0.274*
	(0.072)	(0.040)	(0.048)	(0.086)	(0.156)
Random Part					
σ^2_{u0}	0.211	0.338	0.240	0.466	0.991
	(0.087)	(0.048)	(0.046)	(0.131)	(0.468)
$\hat{R}^2$					
Obs. (level 1)	5.526	29.247	10.096	4.279	1.474
Obs. (level 2)	339	426	371	358	228

Source: Own calculations.
Notes: *P-value<0.1. **P-value<0.01. σ^2_{u0} refers to the variance of the residual errors of the intercepts at the household level (level 2). *** Distance to health facility is measured in kilometers. In the case of Bangladesh distance is measured in time (hours).

Table 4a
Multilevel-Regression of stunting
(full model)

	Bangladesh 2000	India 1999	Mali 2001	Uganda 1995	Zimbabwe 1994
Fixed Part					
Constant	-0.418**	-0.248**	-0.665**	-0.874**	-1.378**
	(0.045)	(0.026)	(0.094)	(0.086)	(0.096)
Age of mother	0.019	0.047*	0.169*	0.050	-0.018
	(0.043)	(0.027)	(0.072)	(0.055)	(0.081)
(Age of mother)^2	0.006	0.009	-0.201*	-0.064	0.046
	(0.073)	(0.045)	(0.112)	(0.087)	(0.124)
Sex of child	0.072	-0.013	-0.208*	-0.279**	-0.151
	(0.069)	(0.035)	(0.119)	(0.084)	(0.122)
Breastfeeding	-0.094	-0.001	0.011	0.044	-0.052
	(0.073)	(0.043)	(0.125)	(0.086)	(0.125)
First born	-0.183	-0.667**	-0.457*	-0.440*	-0.458*
	(0.143)	(0.069)	(0.249)	(0.174)	(0.261)
Preceding birth interval	-0.012**	-0.016**	-0.014**	-0.013**	-0.007*
	(0.002)	(0.001)	(0.004)	(0.003)	(0.004)
Household size	0.017	-0.003	0.004	0.012	-0.014
	(0.015)	(0.005)	(0.008)	(0.020)	(0.023)
Total no. children	-0.002	-0.093**	-0.101*	-0.017	-0.003
	(0.036)	(0.019)	(0.049)	(0.037)	(0.061)
Asset index (global)	-0.341**	-0.114**	-0.107	-0.346**	-0.049
	(0.066)	(0.020)	(0.082)	(0.084)	(0.078)
Mother's education (years)	-0.081**	-0.059**	-0.064*	-0.069**	-0.064**
	(0.013)	(0.005)	(0.035)	(0.016)	(0.023)
Age at marriage	-0.050**	-0.079**	-0.065**	-0.016	-0.017
	(0.016)	(0.009)	(0.024)	(0.0179	(0.025)
Mother's BMI<18.5	0.179*	0.197**	0.234	0.227	0.613*
	(0.072)	(0.037)	(0.202)	(0.151)	(0.246)
Health facility index (global)	-0.134*	-0.090**	-0.266**	0.066	0.024
	(0.054)	(0.023)	(0.079)	(0.069)	(0.089)
Community characterisitcs					
Distance to health facility***	0.001	-0.004*	-0.002	0.004	0.000
	(0.001)	(0.003)	(0.005)	(0.004)	(0.003)
Percent children with fever	0.132	-0.497*	-0.057	-0.492*	-0.096
	(0.220)	(0.240)	(0.360)	(0.206)	(0.290)
Percent secondary education	0.192	-1.075**	-2.726**	-0.857*	-0.103
	(0.264)	(0.189)	(1.016)	(0.338)	(0.340)
Public infrastruct. index	0.022	-0.001	0.018	-0.010	-0.070
	(0.041)	(0.025)	(0.074)	(0.059)	(0.078)
σ^2_{u0}	0.020	0.106	0.000	0.079	0.000
	(0.033)	(0.018)	(0.000)	(0.052)	(0.000)
$\hat{R}^2$					
Obs. (level 1)	3.826	15.180	1.290	2.710	1.633
Obs. (level 2)	339	424	361	357	229

Source: Own calculations.
Notes: *P-value<0.1. **P-value<0.01. σ^2_{u0} refers to the variance of the residual errors of the intercepts at the household level (level 2).*** Distance to health facility is measured in kilometers. In the case of Bangladesh distance is measured in time (hours).

Table 4b
Multilevel-Regression of stunting (new reference standard)
(full model)

	Bangladesh 2000	India 1999	Mali 2001	Uganda 1995	Zimbabwe 1994
Fixed Part					
Constant	-0.129**	0.060*	-0.491**	-0.446**	-1.009**
	(0.044)	(0.026)	(0.094)	(0.079)	(0.093)
Age of mother	0.009	0.066*	0.200**	-0.001	0.003
	(0.042)	(0.026)	(0.070)	(0.053)	(0.075)
(Age of mother)^2	0.015	-0.009	-0.228*	-0.002	0.016
	(0.071)	(0.044)	(0.108)	(0.083)	(0.117)
Sex of child	0.061	-0.131**	-0.111	-0.370**	-0.285*
	(0.067)	(0.033)	(0.117)	(0.080)	(0.113)
Breastfeeding	-0.043	-0.004	-0.180	0.138*	-0.009
	(0.071)	(0.041)	(0.123)	(0.082)	(0.115)
First born	-0.047	-0.523**	-0.326	-0.458**	-0.555*
	(0.139)	(0.065)	(0.243)	(0.163)	(0.236)
Preceding birth interval	-0.009**	-0.014**	-0.016**	-0.011**	-0.008*
	(0.002)	(0.001)	(0.004)	(0.003)	(0.003)
Household size	0.006	-0.006	0.005	0.011	-0.016
	(0.014)	(0.005)	(0.008)	(0.019)	(0.021)
Total no. children	0.003	-0.101**	-0.143**	0.001	-0.041
	(0.035)	(0.018)	(0.046)	(0.036)	(0.057)
Asset index (global)	-0.337**	-0.107**	-0.153*	-0.280**	-0.051
	(0.063)	(0.019)	(0.087)	(0.081)	(0.073)
Mother's education (years)	-0.065**	-0.051**	-0.060*	-0.057**	-0.068**
	(0.012)	(0.005)	(0.034)	(0.015)	(0.021)
Age at marriage	-0.042**	-0.080**	-0.055*	-0.018	-0.029
	(0.015)	(0.009)	(0.023)	(0.016)	(0.023)
Mother's BMI$<$18.5	0.042	0.155**	0.295	0.388**	0.381
	(0.070)	(0.035)	(0.199)	(0.147)	(0.246)
Health facility index (global)	-0.135*	-0.096**	-0.311**	0.015	0.024
	(0.053)	(0.023)	(0.081)	(0.065)	(0.084)
Community characterisitcs					
Distance to health facility***	0.000	-0.008**	-0.001	0.012**	0.002
	(0.001)	(0.003)	(0.005)	(0.004)	(0.003)
Percent children with fever	0.207	-0.360	-1.018**	-0.585**	-0.112
	(0.216)	(0.246)	(0.369)	(0.194)	(0.288)
Percent secondary education	-0.240	-0.970**	-1.681*	-0.878**	0.117
	(0.259)	(0.187)	(0.890)	(0.313)	(0.329)
Public infrastruct. index	-0.028	0.000	-0.061	-0.091	-0.088
	(0.040)	(0.025)	(0.075)	(0.056)	(0.077)
Random Part					
σ^2_{u0}	0.031	0.136	0.085	0.058	0.125
	(0.032)	(0.020)	(0.082)	(0.046)	(0.079)
$\hat{R}^2$					
Obs. (level 1)	3.970	16.540	1.361	2.849	1.683
Obs. (level 2)	339	424	368	356	229

Source: Own calculations.
Notes: *P-value<0.1. **P-value<0.01. σ^2_{u0} refers to the variance of the residual errors of the intercepts at the household level (level 2).*** Distance to health facility is measured in kilometers. In the case of Bangladesh distance is measured in time (hours).

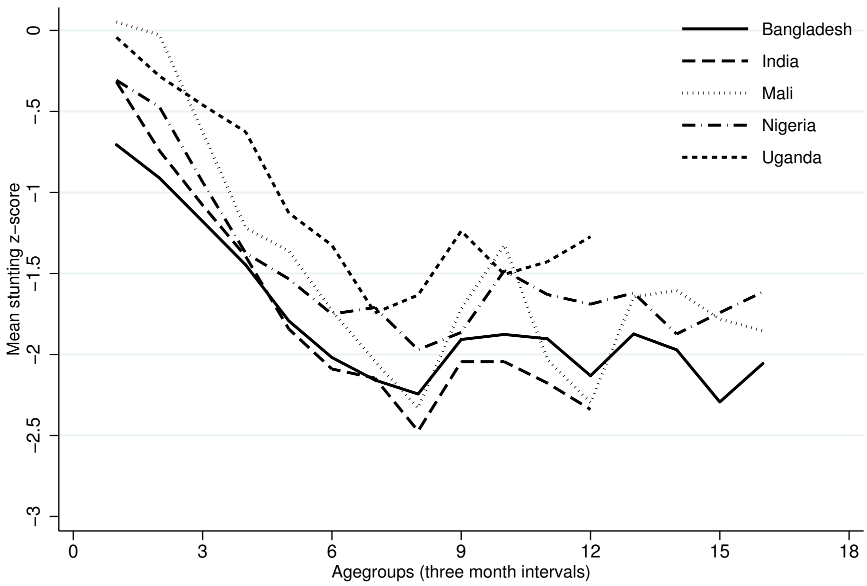
Table 5
Global Regression of child mortality and stunting
(with region fixed and interaction effects)

Constant	-2.796**	(0.046)	-0.167**	(0.031)
Age of mother	-0.247**	(0.020)	0.073**	(0.020)
(Age of mother)^2	0.281**	(0.031)	-0.060**	(0.020)
Breastfeeding	-0.512**	(0.054)	-0.056*	(0.023)
First born	0.521**	(0.076)	-0.355**	(0.038)
Preceding birth interval	0.004*	(0.001)	-0.012**	(0.001)
Household size	0.050**	(0.005)	0.002	0.003
Total no. of children	0.187**	(0.017)	-0.040**	0.010
Asset index (global)	0.051*	(0.024)	-0.097**	(0.011)
Mother's education (years)	0.008	(0.007)	-0.058**	(0.003)
Age at marriage	0.071**	(0.009)	-0.069**	(0.005)
Mother's BMI<18.5	-0.094*	(0.043)	0.208**	(0.021)
Health facility index (global)	-0.836**	(0.026)	-0.127**	0.013
Percent sec. education in cluster	0.358*	(0.168)	-0.658**	0.081
Percent children with fever	0.535*	(0.214)	-0.057	0.107
Region fixed effects				
Sub-Saharan Africa	0.672**	(0.070)	-0.669**	(0.054)
Region interaction effects				
SSA * Breastfeeding	0.360**	(0.073)	0.020	(0.041)
SSA * First Born	-0.318**	(0.116)	0.144*	(0.076)
SSA * Preceding birth interval	-0.007**	(0.002)	0.002	(0.001)
SSA * Household size	-0.041**	(0.006)	-0.004	(0.003)
SSA * Total no. children	-0.018	(0.016)	-0.025	(0.010)
SSA * Asset index (global)	-0.010	(0.041)	-0.026	(0.026)
SSA * Mother's education	-0.027*	(0.014)	-0.004	(0.008)
SSA * Age at marriage	-0.022*	(0.011)	0.029**	(0.007)
SSA * Mother's BMI	-0.063	(0.100)	0.108*	(0.063)
SSA * Health Facility Index	0.204**	(0.045)	-0.020	(0.026)
SSA * Percent secondary education	0.068	(0.287)	-0.364*	(0.175)
SSA * percent children with fever	0.072	(0.278)	0.234	(0.167)
Pseudo R^2				
Obs.	52.991		58.995	

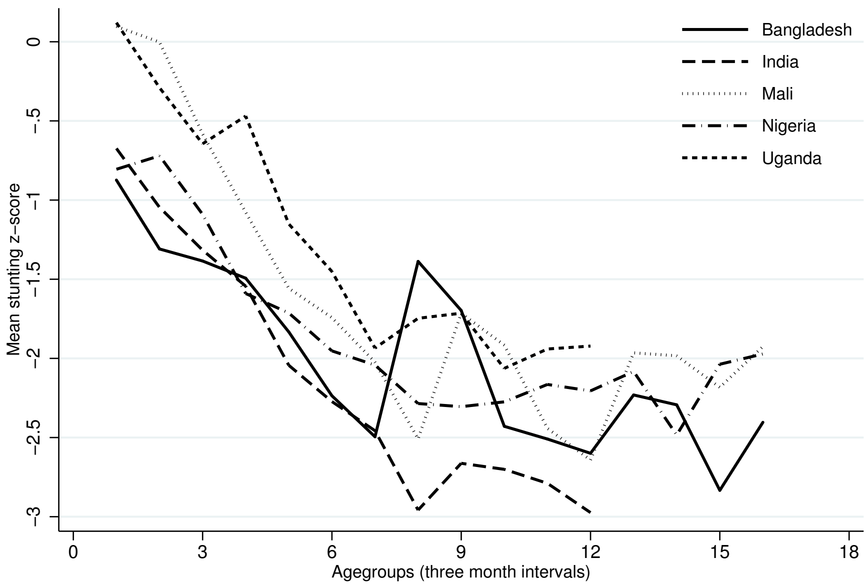
Source: Own calculations.
Notes: *P-value<0.1. **P-value<0.01.

Appendix

Figure A1
(a)
Mean stunting z-score by age



(b)
Mean stunting z-score (new reference standard) by age



Source: Ruban & Ruban calculations.

Table A1
Infant mortality by asset index and health access index quintiles
(percentage)

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Ratio 1/5
			Asset index			
Infant mortality						
Bangladesh 2000	9.26	8.21	8.03	8.12	6.33	1.46
India 1999	10.32	8.43	7.64	7.20	5.69	1.81
Mali 2001	16.54	16.41	15.12	14.41	11.92	1.39
Uganda 1995	11.27	11.16	9.86	9.21	8.46	1.33
Zimbabwe 1994	8.98	7.74	6.50	6.50	8.05	1.12
			Access to health facilities index			
Infant mortality						
Bangladesh 2000	21.72	7.25	7.51	0.81	2.24	9.70
India 1999	15.16	10.80	8.97	3.80	0	n.c.
Mali 2001	21.59	17.29	16.07	11.15	8.08	2.67
Uganda 1995	19.98	16.99	11.12	1.07	0	n.c.
Zimbabwe 1994	23.84	6.52	6.50	0	0	n.c.

Source: Own calculations.

Table A2
Stunting by asset index and health access index quintiles
(percentage)

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Ratio 1/5
			Asset index			
Stunting						
Bangladesh 2000	54.49	49.46	43.01	32.81	16.74	3.26
India 1999	53.11	47.48	44.21	41.16	33.15	1.60
Mali 2001	46.84	40.70	40.62	37.20	23.83	1.97
Uganda 1995	41.57	37.98	31.10	21.77	16.48	2.52
Zimbabwe 1994	22.13	28.12	25.40	20.90	12.43	1.78
			Access to health facilities index			
Stunting						
Bangladesh 2000	56.62	45.90	46.73	31.90	17.51	3.23
India 1999	56.21	54.31	44.64	37.66	31.89	1.76
Mali 2001	43.06	39.85	35.21	21.27	17.16	2.51
Uganda 1995	36.96	39.54	36.32	34.36	33.94	1.09
Zimbabwe 1994	27.71	26.60	21.61	22.52	21.25	1.30

Source: Own calculations.

Table A3
Scoring coefficients for asset index and access to health facilities index
(principal components analysis)

	Bangla-desh 2000	India 1999	Mali 2001	Uganda 1995	Zim-babwe 1994	Global value
Asset index						
Radio	0.191	0.173	0.135	0.173	0.141	0.221
TV	0.284	0.270	0.272	0.245	0.195	0.332
Fridge	–	0.239	0.249	0.182	0.167	–
Bike	0.093	0.077	0.021	-0.002	0.036	0.095
Motorized transport	0.143	0.229	0.205	0.177	0.128	0.263
Low floor material	-0.300	–	-0.255	-0.274	-0.184	–
No toilet facility	-0.125	-0.265	-0.144	-0.118	-0.172	-0.220
Flush toilet	0.273	0.282	0.105	0.195	0.221	0.308
Piped drinking water	0.192	0.196	0.206	0.243	0.203	0.268
Surface drinking water	-0.048	-0.070	-0.085	-0.143	-0.086	-0.142
Access to health index						
Tetanus vaccination	0.393	0.349	0.344	0.480	0.403	0.358
Prenatal care	0.450	0.367	0.347	0.487	0.442	0.376
Born w/o assistance	-0.357	-0.312	-0.335	-0.252	-0.307	-0.286
Vaccinations*	0.303	0.301	0.321	0.334	0.270	0.314

Source: Own calculations.
Note: *Average number of vaccinations per child in respective age in household.