University of Innsbruck

Multilevel modelling of child mortality in Africa

Kenneth Harttgen, Stefan Lang, Judith Santer

Working Papers in Economics and Statistics

2015-03

Research Platform eeecon
Empirical and Experimental Economics
University of Innsbruck
http://eeecon.uibk.ac.at/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
Research platform “Empirical and Experimental Economics”
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 7171
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
http://eeecon.uibk.ac.at/wopec/

For a list of recent papers see the backpages of this paper.

Multilevel Modelling of Child Mortality in Africa

Kenneth Harttgen[a], Stefan Lang[b], and Judith Santer[b]

[a] *NADEL, ETH Zürich, Clausiusstr. 37, 8092 Zurich, Switzerland*

[b] *Department of Statistics, University of Innsbruck, Universitätsstr. 15, 6020 Innsbruck, Austria*

Abstract

Whereas child mortality has been decreased globally in the last 20 years, high levels persist in Sub-Saharan Africa. This paper analyzes child mortality in 25 sub-Saharan countries based on household survey data. We employ a new multilevel approach with structured additive predictor within the Bayesian framework. This allows us to take into account the hierarchical data structure and use the heterogeneity within and between countries as well as to assess non-linearities in the relationship between child mortality and socio-economic determinants. We find that household's economic well-being, mother's education and age, and geographical regions strongly influence child mortality risks.

Keywords: Child mortality, sub-Saharan Africa, multilevel STAR models, Bayesian inference.

JEL-Codes: C11, I12

1 Introduction

Child mortality varies among world regions but the highest prevalence is concentrated in Sub-Saharan Africa where mortality of children under five decreased from 177 in 1990 to

98 deaths per 1,000 live births in 2012 (UNICEF, 2013). Despite the overall decline in the prevalence of child mortality, it remains still at unacceptably high levels. About half of all deaths of children under five has been concentrated in Sub-Saharan Africa in 2012 (UNICEF, 2013). Hence, the need to reduce child mortality is one of the major challenges in improving child health, in particular in Sub-Saharan Africa.

Child health affects health in adulthood and socioeconomic status. Investments in child health made during critical periods of child development result in larger returns; conversely, failure to invest can lead to severe long-term economic consequences at both the household and the macro-economic level. A growing evidence of the long term benefits of investing in children exits highlighting their importance to improve the well-being of the child (e.g. Almond and Currie, 2010; Bhalotra and Rawlings, 2011; Case et al., 2005; Currie, 2009). Causes of deaths differ substantially from one country to another and within countries between regions, which highlight the need to expand understanding of the determinants of child health at a country level (and within countries between regions) rather than simply focusing on geopolitical regions.

Analyzing the socio-economic determinants of child mortality is not new. A broad body of literature has investigated the determinants of child mortality in developing countries during the last 30 years. Generally, economic research analyses the underlying causes of child health outcomes by evaluating constraints in individual's opportunities related to child health (Schultz (1984); Mosley and Chen (1984)). The mechanisms of how differences in individual endowments affect child health outcomes are complex and factors influencing child health outcomes differ across regions, countries and within countries across population subgroups. Although no distinct factors can be identified that are generally applicable to all situations, summarizing the existing empirical literature reveals a general pattern of underlying factors of child health in developing countries.

At the individual and household level the socioeconomic status as well as environmental factors are identified to be important determinants of child health outcomes. Several single-country studies based on micro data have shown the impact of individual's or household's endowments of resources (e.g. income, assets, land) as well as access to safe drinking water, food, energy, and improved sanitation on infant and child mortality (Kembo and Van Ginneken (2009) (Zimbabwe); Mesike and Mojekwu (2012) (Nigeria); Gemperli et al. (2004) (Mali); Nuwaha et al. (2011) (Uganda); Manda (1999) (Malawi); Kandala and Ghilagaber (2006) (Malawi); Adeyemi et al. (2008) (Nigeria); Adebayo and Fahrmeir (2005) (Nigeria); Ogunjuyigbe (2004) (Nigeria); Wang (2003) (Ethiopia). For example, using demographic and health survey (DHS) data from Kenya, Mutunga (2007) shows strong positive impacts of socioeconomic and environmental factors on child survival. He also investigates the relative importance of socioeconomic endowments and environmental factors by the age of the child. While birth spacing and breastfeeding are found to be relatively more important for the survival probability during the period of infancy, socioeconomic variables and environmental factors such as access to save water, improved sanitation, or indoor air pollution are relatively more important with increasing age of child. Similar results are found by Kyei (2012) for South Africa. The effect of maternal factors such as the age of the mother is found to be limited to infant mortality, while socioeconomic variables are found to be more important with increasing age of child. Also, Kembo and Van Ginneken (2009) show for Zimbabwe that the socioeconomic status and access to sanitation is relatively more important for child survival whereas increases of birth spacing and accessibility to health infrastructure and family planning are relatively more important for infant survival than for child survival.

Maternal and parental education has also found to be strong determinants of child health outcomes. More education is not only associated with higher socioeconomic status (e.g. Caldwell (1979), 1989; Caldwell and McDonald (1982); Caldwell et al. (1983)) but it is also a determinant of treatment factors of illness (e.g. Ojikutu (2008) and Ojikutu (2010)). Furthermore, empirical evidence exists showing the impact of health and population policies on the availability and accessibility of public and private services. For example, Becher et al. (2004) find strong evidence for the importance of the accessibility to health care in Burkina Faso. Also Armstrong-Schellenberg et al. (2002) show for Tanzania that the household environment including accessibility of health care and individual care reduce the risk of child mortality. Similar results are found by Adeyemi et al. (2008) for Nigeria. Other researchers provide cross-country evidence on the importance of socioeconomic and environmental factors and access to public and private health infrastructure across countries using multiple standardized household survey data. For example, Bicego and Boerma (1993) use DHS data for 17 developing countries and find that maternal education has a strong positive impact on child health and survival. Also using DHS data for five developing countries, Harttgen and Misselhorn (2006) find that socioeconomic factors such as the asset ownership of the household as well as the accessibility to community health infrastructure positively influence child and infant survival rates. Using the world fertility surveys (WFS) Hobcraft et al. (1984) find that the education of mothers and husbands are important for child health outcomes linking the education of parents to their occupation and their socioeconomic status. Kazembe et al. (2012) use census data from Rwanda, Senegal, and Uganda and identify environmental factors, urbanization, bio-demographic characteristics, and socioeconomic status as the most important determinants of child survival.

A third strand of empirical research used aggregated macro data to study the determinants of child health outcomes. For example, Hanf et al. (2013) show, based on a longitudinal study for 193 countries using annual data between 2000 and 2009 from the World Development Indicators, that GDP per capita, access to safe drinking water, improved sanitation, and public health expenditure per capita increases the probability of child survival. Pelletier and Frongillo (2003) analyze the effect of undernutrition on child survival using longitudinal data at the national and subnational level between 1966 and 1996. Their results indicate that undernutrition has a strong negative impact on child survival that is independent from socioeconomic and policy changes. Pritchett and Summers (1996) analyze the impact of income of child health outcomes across country and over time for a period between 1980 and 1990. Using an instrumental variables approach to capture the pure income effect they find a long run income elasticity of infant and child mortality between -0.2 and -0.4.

In sum, existing studies have shown that child mortality is influenced by individual, household (e.g. education, wealth), and community or regional conditions (e.g. infrastructure, climate, disease, environment) that directly and indirectly determine proximate or intimidate input variables for child health outcomes (e.g. water and sanitation, hygiene nutrition, medical care (Mosley and Chen (1984); Schultz (1984); Wolpin (1997)). However, evidence on why high levels of child mortality persist in Sub-Sahara is still limited and open and the question is still for debate. Furthermore, no study exists that analyzes the association between child mortality and socioeconomic characteristics using various household survey data from several countries and years simultaneously, taking into account the clustered data structure of the survey data and analyzing non-linearities in the relationship between child mortality and socioeconomic characteristics.

In this paper, we analyze child mortality as an (reverse) indicator of child health. The main objective of our study is a better understanding of underlying causes of child mortality and variations in child health outcomes in Sub-Saharan Africa. Two features of the analysis in this study contribute to a better understanding of determinants of child health. First, the study uses the largest available, nationally representative, and mutually comparable repeated cross-sectional samples on as many as 315,721 children born to 91,688 mothers in 25 countries from 62 surveys in the period of 1992 to 2009. Using a pooled large-scale household survey sample provides us with exceptional heterogeneity within and between countries to analyze the association between child mortality and socio-economic characteristics. In addition, this allows us to analyze general patterns in the determinants of child mortality in Sub-Saharan Africa as well as to make inter-regional comparison within countries. The focus on Sub-Saharan Africa is hereby relevant because child health is still at the lowest level in Sub-Saharan Africa compared to all other regions of the world. Second, we employ a multilevel approach with structured additive predictor within a Bayesian framework. This approach has two important advantages. First, the multilevel approach takes into account the hierarchical structure of the household survey data. Children/households (level-1) are nested within regions (level-2) and regions are in turn nested within countries (level-3). The multilevel approach explicitly takes into account this data structure correcting for the bias in the parameter estimates resulting from the nested data structure, because each level is represented by its own sub-model, which expresses the association between child mortality and explanatory variables within that level. Moreover, regional heterogeneity in covariate effects is explicitly modeled and uncovered within the multilevel approach. Second, the modeling framework allows assessing non-linearities in the relationship among child mortality and socio-economic characteristics. The estimation of our models is done by using the R package **BayesR** (Umlauf et al., 2013b).

The rest of the paper is organized as follows: Section 2 provides a description of the data and of the conceptual framework for our empirical approach as well as descriptive statistics. Section 3 is devoted to the representation of the modelling framework, Section 4 shows the results and finally, Section 5 concludes.

2 Data and Descriptive Analysis

We use nationally representative Demographic and Health Surveys (DHS) for this study. DHS are standardized surveys that have been undertaken by Macro International Inc., Calverton, Maryland (usually in cooperation with local authorities and funded by USAID). Established in 1984, more than 200 standard DHS from around 70 countries have been conducted. For most countries more than one survey (up to seven) is available. The average sample size is about 5,000 to 30,000 households. The DHS provide detailed information of women, men, children and households on population, health, nutrition, and education as well as on household's durables and quality of the dwelling. Within each sampled household the household roster was taken and women eligible for a more detailed interview were identified. For each eligible woman between 15 and 49 birth histories were recorded. From these birth histories, information on child mortality and age at death are provided. Our sample covers all African countries. The study sample consists of 62 surveys from

25 countries resulting in information on 315,721 children born at least five years before the survey. A detailed list of the number of children per country and survey year can be found in the appendix A.1. Figure 1 shows a map of Africa with grey coloured areas representing the countries being part of our analysis. These areas show the countries in the study sample, which cover two regions in Sub-Saharan Africa: western and eastern Sub-Saharan Africa. In addition, the DHS present a selected sample of more successful African economies, because countries that are in serious economic troubles, or even face civil conflict, are unlikely to be able to field a DHS. Countries within these two regions show within and between country variations in mortality and socioeconomic characteristics reflecting no structural differences between these two geographical regions that could bias our results.

[INSERT FIGURE 1 ABOUT HERE]

The underlying theoretical framework for the choice of the dependent and independent variables to study child mortality and undernutrition, i.e. classification of child, parent and household characteristics into socioeconomic and demographic factors, closely follows the analytical framework proposed by Mosley and Chen (1984) to study child survival. The idea of this framework is the assumption that social, economic, demographic, and medical determinants, i.e. the proximate determinants, affect the survival probability of the children through a set of biological mechanisms (see also Sullivan et al., 1994; Majumder et al., 1997). In this analysis, the Mosley and Chen (1984) framework is combined with the conceptual framework to study the causes of child undernutrition proposed by the United Nations Childrens' Fund (UNICEF, 1990), which is based on assumptions similar to the Mosley and Chen (1984) framework, and the subsequent extended model of Engle et al. (1999), which implements also the provision of health care capacities of households into the analysis of childrens' welfare. Below, we will describe the variables used in our case study as proximate determinants of child mortality, which are grouped at different hierarchical levels, i.e. the individual/household, region, and country level.

2.1 Explanatory Variables

Demographic factors

Measure for malnutrition of children and mothers

In this study, we focus on the anthropometric indicator height-for-age (stunting), which provides information on chronic undernutrition at the time of interview. Besides height-for-age, weight-for-age (underweight) and weight-for-height (wasted) are commonly used anthropometric indicators. The value of the index for child i is related to a reference population by median subtraction and standard deviation division, i.e. defining a Z-score

$$Z\text{-}score_i = \frac{ai_i - ai_{med;r}}{\sigma_r}$$

with ai_i representing the anthropometric index of child i, $ai_{med;r}$ indicating the median anthropometric index and σ_r the standard deviation of the reference population. For stunting, the Z-score is calculated by using the child's height minus the median height for that child's age and sex in a reference population divided by the standard deviation of this group in the World Health Organization reference population of health children in developing countries (WHO, 2006).[1] A child is then defined as stunted if the height for age Z-score is less than minus two relative to children of the same sex and age in the reference population.[2] A child is defined as severely stunted if the height for age Z-score is less than minus three. Unfortunately, we cannot analyze the impact of undernutrition at the individual (child) level, because the data provides no information on anthropometric indicators of children who have died. Therefore, we measure the impact of undernutrition on child mortality at the next higher level by aggregating the means in stunting for each geographical region within countries.

In order to measure the nutritional situation of adults another well known index has been included in the analysis. The Body-Mass-Index (BMI) relates the weight of adult i (measured in kilograms) to the squared height of adult i (measured in metres). The thresholds for underweight, normal weight and overweight are <18.5, 18.5-24.99 and ≥ 25, respectively (see WHO, 2000).

Additional child and mother characteristics

The gender of the child as well as the parity order (first born child) are taken into account. Furthermore, we are interested if deaths of siblings in the past have an influence on mortality under five. Besides the BMI of the mother, her age at the time of the interview is of interest.

Socioeconomic factors

Household size

The household size generally includes the number of children (≤ 18 years) and the number of adults (> 18 years). As already mentioned above, the age of the mother will be included in the model. Since elder women might have more children, we exclude the number of children living in the respective household and instead represent the household size by the number of adults.

Literacy

An early reference for analyzing the effect of maternal education on child mortality is Caldwell (1979) who used Nigerian survey data. Since mothers are, in general, responsible for child-rearing, the educational level of the mother is important for the well-being of the child. The argument here is twofold. First, more educated women might be able to better process information and to acquire skills in order to take care of the children, for example in the case of illness, and second, better educated women are in a better position to earn money. Several studies have shown evidence that children's health status benefits from mothers who are at least able to read (e.g. Blunch, 2013; Smith-Greenaway, 2013). Thus, we include the variable *mliterate* in the analysis. In accordance with the DHS, a literate person has attended secondary or higher education or is able to read a whole sentence.

Place of residence

This variable indicates where the household is located, either in urban areas (cities, towns) or in rural areas. In addition, we assume that the various effects might differ across regions or countries and thus include the variables *region* and *country* that assign the different observations to a particular region and country. This (geographical) classification of observations into groups is called clustering (see e.g. Jain et al., 1999). Accordingly, a particular region cluster contains all observations that pertain to this specific region.

Measures for wealth

I.) Measure at individual level - Asset index

Unfortunately, household income or expenditures are not included, but available information on household assets can be used to construct an asset index proxying the household's long-term well-being. In order to generate a measure for wealth on the individual (household) level to overcome the problem of missing income and expenditure data, Filmer and Pritchett (2001) and Sahn and Stifel (2003) have proposed a one-dimensional index based on household assets and other household characteristics as a proxy of long-term material welfare. The proposed asset index has been widely used in the empirical literature on poverty and inequality analysis as a proxy variable for the material welfare of the household.[3] We follow this approach and include the following assets: electricity, radio, television, refrigerator, bicycle, motorbike, car, low floor material, no toilet facility, source of drinking. For the estimation of the weights and for the aggregation of the index, we use a principal component analysis as proposed by Filmer and Pritchett (2001), where the first principal component is the asset index.

II.) Measure at country level - Real gross domestic product per capita

The individual data from the Demographic and Health Surveys are merged with the real gross domestic product per capita at 2005 constant prices from the Penn World Table 7.0 (2012).[4] The unit of observation is the child, and children in the same survey (same country and year) have the same GDP per capita which is representative of the national average. We merged the Demographic and Health Surveys over survey year the DHS has been conducted.

Coding and Description

A list of variables with information concerning the coding, the hierarchy level as well as the mean and relative frequency of the continuous and categorical variables can be found in Table 1.

[INSERT TABLE 1 ABOUT HERE]

Child mortality is the variable of primary interest and yields information on the survival status, e.g. dead or alive, of children who are born at least five years before the survey. The response will be dummy coded, i.e.

$$\text{dead5} = \begin{cases} 1, & \text{if the child is dead} \\ 0, & \text{otherwise.} \end{cases}$$

The remaining categorical covariates will be effect coded mainly for technical reasons. The speed of convergence and the mixing of sampled parameters in Markov chain Monte Carlo based Bayesian estimation of the models is improved compared to dummy coded covariates (Umlauf et al., 2013a). Consider a covariate x with e.g. three categories of whom category

2 represents the reference category. Defining an effect coded variable $x_{(1)}$ is done as follows

$$x_{(1)} = \begin{cases} 1, & \text{if x=1} \\ -1, & \text{if x=2} \\ 0, & \text{if x=3.} \end{cases}$$

The variable $x_{(3)}$ for category 3 is defined analogously. In contrast to dummy coding, effect coding produces new variables that are coded with -1 for the reference category yielding a sum to zero constraint (Fahrmeir et al., 2013). In case of only two distinct values, the effect coded variable has the values 1 and -1 (see e.g. the variable *gender* in Table 1).

3 Statistical Modelling

3.1 Binary regression with linear and additive predictors

Since the variable of primary interest *dead5* is binary, a natural class of regression models are logit or probit models assuming (conditional on covariates) a Bernoulli distribution, i.e. $dead5_i \sim B(1, \pi_i)$, for the i-th observation $dead5_i$ with parameter or probability $\pi_i = P(dead5_i = 1) = E(dead5_i)$. For a probit model as is assumed in this paper we obtain

$$\pi_i = \phi(\eta_i),$$

where $\phi(\cdot)$ is the cumulative distribution function of the standard normal distribution and η_i is a suitable predictor that models the dependence of $dead5_i$ on the covariates. Traditionally, a linear predictor is assumed, i.e.

$$\eta_i = \eta_i^{lin} = \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \cdots + \gamma_p x_{pi}. \quad (1)$$

Here, $x_{1i}, \ldots, x_{pi}$ are the covariates. A positive regression coefficient γ_j indicates that the predictor η and thereby also the probability π increases as x_j increases. However, the increase is nonlinear because of the nonlinear cdf ϕ.

In most case studies the assumption of a strictly linear predictor (1) is not appropriate because of possibly nonlinear relationships for the continuous covariates. This restriction can be solved in modelling a continuous covariate effect by a possibly nonlinear function f as in

$$\eta_i = \gamma_0 + \gamma_1 \, x_{1i} + f(x_{2i}) + \ldots,$$

where the effect of x_2 is assumed to be nonlinear with function f and all other effects still being linear. If the effects of all covariates are assumed to be nonlinear while the additive decomposition of covariate effects is still pertained the model is a so-called additive model. In our case, a possible predictor is given by

$$\begin{aligned}\eta_i \quad = \quad & \gamma_0 + f_1(agemother_i) + f_2(adults_i) + f_3(mbmi_i) + f_4(ai\_drm_i) + \\ & \gamma_1 gender_i + \gamma_2 fhh_i + \gamma_3 firstbchild_i + \gamma_4 deadsibling_i + \gamma_5 mliterate_i + \gamma_6 urban_i + \\ & \gamma_7 wave2_i + \gamma_8 wave3_i + \gamma_9 wave5_i + \gamma_{10} wave6_i. \end{aligned} \tag{2}$$

The continuous covariates *agemother, mbmi, adults* and *ai_drm* are assumed to affect the response in a nonlinear way, whereas for the effect coded categorical covariates *gender*, *fhh*, *firstbchild*, *deadsibling*, *mliterate*, *urban* and *wave* linear relationships are kept. The time component is represented by the covariates $wave_t$, that refer to the official survey wave of the DHS and that provide the information about the period when the survey was conducted with $t \in \{2, 3, 5, 6\}$.[5] We do not use year fixed effects (the year the survey was conducted) to take into account that countries are not equally distributed across time and that each country has only been surveyed once during each DHS wave. Due to the fact, that for a majority of countries the survey has been conducted at the $4th$ wave, we use the period from 1998 to 2005 ($=wave_4$) as reference category.

[INSERT FIGURE 2 ABOUT HERE]

The interpretation of a nonlinear relationship is best undertaken by plotting the estimated function against the covariate values. This is done exemplarily for the nonlinear function f_1 in Figure 2. The function is almost linearly decreasing until the age of 33 years indicating that the probability of dying decreases for children with the age of the mother until the threshold of 33 years is reached. Thereafter, the effect is almost constant and thereby also the probability of dying. The confidence intervals in Figure 2 can be seen as a measure of uncertainty of the estimate. For instance, the uncertainty is higher at the left border of the covariate domain.

Technically, a nonlinear function f is approximated by a linear combination of appropriate basis functions, i.e.

$$f(x_i) = \beta_1 \, B_1(x_i) + \beta_2 \, B_2(x_i) + \ldots + \beta_l \, B_l(x_i). \tag{3}$$

In the three panels of Figure 3 we illustrate the approximation for some artificial data. First, the range of the covariate is divided in equally spaced intervals. Then the basis functions are constructed such that within each interval f forms an arbitrary polynomial of degree three and the polynomials are continuously connected[6]. The basis functions are depicted in the left panel of Figure 3. Next, the unknown coefficients $\beta_1, \ldots, \beta_l$ are estimated. Multiplying each coefficient with its respective basis yields the scaled basis functions in the second panel. Finally, adding the scaled basis functions leads to the estimated effect in the right panel.

[INSERT FIGURE 3 ABOUT HERE]

The approximation (3) transforms the nonlinear predictors into linear ones. In principle the models could then be estimated using standard maximum likelihood methods for binary regression models (see Fahrmeir et al., 2013, Chapter 5). However, the choice of the number of knots (or intervals) to approximate the unknown nonlinear functions is a delicate issue in practise. A small number of intervals gives a tendency to underestimate the complexity of some curves whereas a large number of knots often leads to relatively wiggly estimates that are difficult to interpret. Therefore, we use here a penalized likelihood approach described in full detail in Fahrmeir et al. (2013), Chapters 8 and 9. Basically, the main idea is to define a relatively large number of knots (the default choice is 20 knots) to prevent underfitting and at the same time to introduce a roughness penalty for the regression coefficients to prevent overfitting. Details are beyond the scope of this paper and can be found in the reference provided above.

3.2 Multilevel structured additive regression

As already mentioned in the introduction, a key aspect of our analysis is the hierarchical structure of our data: 315,721 children are nested within 324 regions and these regions are in turn nested within 25 countries. This gives rise to a hierarchical or multilevel version of our additive models. Adopting the hierarchical data structure to our application yields the following two level predictor with individual level covariate effects at the lowest level of the hierarchy and region specific explanatory variables at level-2:

$$\begin{array}{lrcl} \text{level-1:} & \eta_i & = & f_1(agemother_i) + f_2(mbmi_i) + f_3(adults_i) + f_4(ai\_drm_i) \\ & & & + f_5(region_i) + x_i^\top \gamma \\ \text{level-2:} & f_5(region) & = & f_{5,1}(stunting\_rm_{region}) + f_{5,2}(country_{region}) \end{array}$$

The effect coded categorical covariates in (2) are now subsumed in the term $x_i^\top \gamma$. Besides linear effects and the continuous covariate effects of *agemother, mbmi, adults* and *ai_drm* the level-1 equation now additionally contains an i.i.d. Gaussian random effect $f_5(region) \sim N(0, \tau_R^2)$ to cope with unexplained region specific heterogeneity. The region specific random effect is further decomposed in the level-2 equation. The nutritional situation of children is quantified by the Z-score based on the anthropometric indicator height-for-age, i.e. stunting. Data on stunting is only available for children still being alive. Thus, we include the average Z-score per region at the second level to model an external effect. We control for country specific differences by incorporating another i.i.d Gaussian random effect in the level-2 equation, denoted by $f_{5,2}(country_{region})$. Table 2 gives an overview of the expected effects of the covariates on the response variable.

[INSERT TABLE 2 ABOUT HERE]

The model can be further extended by introducing additional country specific random slope effects for some covariates. Considering all covariates listed in Table 1 leads to the following full model:

$$\begin{array}{llcl}
\text{level-1:} & \eta_i & = & f_1(agemother_i) + f_2(mbmi_i) + f_3(adults_i) + f_4(ai_drm_i) \\
& & & + f_5\,(region_i) + f_6\,(country)\,gender_i + f_7\,(country)\,fhh_i \\
& & & + f_8\,(country)\,firstbchild_i + f_9\,(country)\,deadsibling_i \\
& & & + f_{10}\,(country)\,mliterate_i + f_{11}\,(country)\,urban_i \\
& & & + f_{12}\,(country)\,wave2_i + f_{13}\,(country)\,wave3_i \\
& & & + f_{14}\,(country)\,wave5_i + f_{15}\,(country)\,wave6_i \\
& & & + f_{16}\,(country)\,agemother_i + f_{17}\,(country)\,adults_i \\
& & & + f_{18}\,(country)\,mbmi_i + f_{19}\,(country)\,ai_drm_i \\
\text{level-2:} & f_5(region) & = & f_{5,1}\,(stunting_rm_{region}) + f_{5,2}\,(country_{region}) \\
\text{level-3:} & f_{5,2}(country) & = & f_{5,2,1}\,(rgdp_{country})
\end{array} \tag{4}$$

Compared to the previous model, model (4) now also includes country-specific random slopes (denoted by the terms $f_k(country)$ for $k = 6, 4, \ldots, 19$) as well as the real gross domestic product per capita as an additional variable for wealth at country level. For model choice we mainly rely on a Bayesian goodness of fit criterion, the deviance information criterion (DIC) introduced by Spiegelhalter et al. (2002). The DIC shows a trade-off between model fit and complexity with an interpretation similar to the well known AIC. When comparing two models, a rule of thumb proposed by Spiegelhalter et al. (2002) indicates a clear preference for one of the two models under consideration if the DIC difference is larger than 10 points whereas it is indecisive otherwise. In the latter case, we assist the model selection by looking at the significances of the estimated effects, i.e. we exclude a parametric effect from the optimal model if the 95% credibility interval contains zero. For nonlinear effects, we base our decision on simultaneous credible bands (Krivobokova et al., 2010) and exclude nonlinear effects if such a 95% credible band contains zero for all observed covariate values. Performing a stepwise selection process we finally end up with the following model:

$$
\begin{array}{lrcl}
\text{level-1:} & \eta_i & = & f_1(agemother_i) + f_2(mbmi_i) + f_3(adults_i) + f_4(ai\_drm_i) \\
 & & & +f_5\,(region_i) + x_i^\top \gamma + f_6\,(country)\,firstbchild_i \\
 & & & +f_7\,(country)\,deadsibling_i + f_8\,(country)\,mliterate_i \\
 & & & +f_9\,(country)\,urban_i + f_{10}\,(country)\,wave2_i \\
 & & & +f_{11}\,(country)\,wave3_i + f_{12}\,(country)\,wave5_i \\
 & & & +f_{13}\,(country)\,wave6_i + f_{14}\,(country)\,agemother_i \\
 & & & +f_{15}\,(country)\,adults_i + f_{16}\,(country)\,mbmi_i \\
 & & & +f_{17}\,(country)\,ai\_drm_i \\
\text{level-2:} & f_5(region) & = & f_{5,1}\,(stunting\_rm_{region}) + f_{5,2}\,(country_{region})
\end{array}
\tag{5}
$$

For comparison we additionally estimated models for each country separately.

For full details on the technical aspects of the multilevel modelling framework applied in this paper see Lang et al. (2014).

4 Results

4.1 Nonlinear and fixed effects

Figure 4 depicts the estimates of nonlinear effects at level-1 and level-2 obtained from Equation (5). Shown is the posterior mean together with 95% credible intervals. Beginning with the largest effect, the age of the mother (variable *agemother*), the probability of dying considerably decreases for children below the age of five with the age of the mother (Ojikutu, 2008). This effect is particularly strong up to the age of the mother of about 33 years. Then, the effect remains constant up to an approximate age of 40 years before the probability slightly increases. For very young women, the probability of facing the death of a child is considerably higher.

[INSERT FIGURE 4 ABOUT HERE]

The effect of the household size (modelled by the number of adults, variable *adults*) seems to be not significant as the 95% credible interval almost fully covers the zero line when controlling for the socioeconomic characteristics of the household.

The most interesting new results are found for the effect of the nutritional status of the mothers. The effect of the BMI of the mother (variable *mbmi*) on child mortality is not linear, but it seems to follow an inverted U-shape (for the most relevant range between 18 and 27), contrary to our expectations. First, the probability increases up to a BMI of 25 and then decreases. However, the inverted U-shape relationship of the BMI of the mother is a frequent result in the analysis of the determinants of undernutrition, see e.g. Kandala et al. (2009) and Bhalotra and Rawlings (2011). This finding indicates that the magnitude of effect of the mother's nutritional status depends on the level of the BMI, when controlling for other socio-economic characteristics. This shows that the depths of undernutrition of women is important indicating that improvements in child survival demand considerable improvements in the nutritional status of women.

The standard of living is proxied by the asset index (variable *ai_drm*) and modelled as an individual-/household-specific deviation from the regional mean. Also the effect of the standard of living of the household on the probability of child mortality reveals some interesting findings. The effect of an asset index below the regional average on child mortality seems to be almost constant. A household whose wealth reaches beyond the regional average first exhibits a considerable decline in the probability of child mortality before the effect slightly levels off. This finding indicates that improvements in the material welfare of the household can reduce child mortality emphasizing the need for policies fostering economic (pro-poor) growth.

Finally, we obtained estimates for the fixed effects of female headed household (variable *fhh*) and sex of the child (variable *gender*). With respect to the latter, the results seem to be in line with our expectations, i.e. male children seem to have a significantly higher risk of dying before the age of five (pmean = 0.0317; 95% credible interval $[0.0263; 0.0368]$) (see also Harttgen and Misselhorn, 2006). For children living in a household with a female head the survival chances are significantly lower as well (pmean = 0.0187; 95% credible interval $[0.0136; 0.0239]$) (see also Kembo and Van Ginneken, 2009).

At the higher hierarchical level the effect of the regional average of child undernutrition rates (variable *stunting_rm*) also show some interesting nonlinear effects. As the regional Z-score increases which means that the chronic malnutrition decreases, the probability of mortality for children significantly declines. Therefore, high levels of undernutrition are associated with higher rates of child mortality (see e.g. Klasen, 1996) indicating a close relationship between the nutritional status of children and their survival probability. This finding also has important policy implications. In particular, investments in the nutritional status of children can also have positive effects on their survival probability.

4.2 Random effects

Figure 5 depicts the random slopes of the continuous and categorical covariates. The first four panels illustrate the effects of the continuous covariates. The respective main effects $f_1 - f_4$ are denoted by black lines. The remaining curves (grey lines) provide the 25 country-specific regression curves (main effects plus random slopes). The effect of *agemother* with the strongest impact on mortality seems to be heterogeneous across countries. Also the effect of the BMI of the mother (variable *mbmi*) varies considerably (note that the effects are not plotted on the same scale). In some countries the slope coefficient is relatively high, i.e. the nonlinear effect on mortality is relatively steep, while in others the effect is nearly zero. This is a very interesting finding with respect to the formulation of policies to reduce child mortality. Although policies such as improvements of the material welfare are important across countries, country specific heterogeneities in the effects of determinants of child mortality need also to be considered.

[INSERT FIGURE 5 ABOUT HERE]

The next four panels illustrate the random effects of the categorical covariates. In doing so, we plot the posterior means together with the 95% credible intervals in ascending order. The credible intervals are defined by the lower and upper grey coloured bars. The effects are assumed to be significant if the credible interval does not contain zero. In case of significance, the lower and upper bars are dark grey coloured, otherwise they are light grey coloured. Furthermore, the country-specific random effects of the categorical covariates differ not only in magnitude and direction but also in significance. Whereas the effects of first born children (variable *firstbchild*), children with a dead sibling (variable *deadsibling*) and place of residence (variable *urban*) seem to be relatively small and not significant in some countries when controlling for the standard of living of the household, the opposite is true considering the effect of the literacy of the mother (variable *mliterate*). The latter influences child mortality negatively, i.e. if the mother has attended secondary or higher

education or is able to read a whole sentence the probability of mortality decreases.
In contrast to our expectations, first born children have a significantly lower risk of dying in ten out of 25 countries. As anticipated deaths of siblings in the past increase the risk of facing mortality under five. In almost two-thirds of the analyzed countries this effect is very similar but not significant. In Liberia, Madagascar and Nigeria the risk of facing mortality is significantly higher for children who have already lost a sibling. The effect of *urban* is ambiguous: in Niger and several other countries children who reside in urban areas are less affected by mortality, whereas in Kenya the opposite seems to be true. The time effects modelled by introducing wave variables into the model and depicted in the following three panels are not that clear. These panels only illustrate the results of countries that have been part of the respective survey wave[7]. During the period from 1990 to 1992, child mortality seems to be higher compared to the reference category, although this effect is not significant. Whereas the risk of child mortality seems to be (significantly) higher in periods previous to our reference category (e.g. in Benin, Burkina Faso, Niger and Ghana during *wave3*), child mortality seems to be (significantly) lower in periods afterwards (e.g. Niger and Ghana). This might reveal the trend in declining mortality rates in some countries.
Finally, the remaining panel depicts the country random intercept estimates that show the country-specific differences in mortality. These intercept coefficients differ considerably in magnitude and variability from the slope coefficients.

5 Conclusion

The persistence of large levels of child mortality in Sub-Saharan Africa remains one of the major concerns in improving health and development in this region. In this paper, we employed a new multilevel approach with structured additive predictor within the Bayesian framework to analyze determinants of child mortality. In particular, this paper analyzed the determinants of child mortality as an (reverse) indicator of child health in using survey data of 315,721 children under five in 25 sub-Saharan African countries. Each observation was assigned to one out of 324 regions in 25 countries. This hierarchical structure and the assumption of possibly nonlinear effects leads to the use of the multilevel structured additive regression approach of Lang et al. (2014).

The selection of variables was primarily motivated by the proposal of Mosley and Chen (1984) on the one hand and by the availability of those explanatory variables in the DHS data.

The use of the multilevel approach and the introduction of a potential nonlinear relationship between child mortality and its socio-economic determinants reveals interesting results that help to better understand why high levels of child mortality exist in Sub-Saharan Africa. First, we find that the living standard of a household plays a crucial role in reducing the risk of mortality. Considering the effect of wealth at household level proxied by an asset index shows evidence of a decline in mortality with increasing wealth. This decline is more pronounced as the household wealth reaches beyond the regional average. Second, male children face a higher risk of dying than their female counterparts. Third, first born children are less affected by mortality. Fourth, the mother's ability to read and write may increase the survival chances of children. Fifth, urban areas exhibit a lower risk of mortality. This is in line with some findings in the literature, e.g. Balk et al. (2003) and van de Poel et al. (2009), who state that children living in urban areas face a lower risk of dying due to e.g. improved infrastructure and sanitation facilities. In contrast, Fay et al. (2005) accuse the pollution and crowding responsible for the higher mortality risk of children living in urban areas. Similar results have been found by Adebayo and Fahrmeir (2005) who analyzed child mortality in Nigeria, although the effect seems to be insignificant. Our results show that the effects of determinants of child mortality are often misleadingly considered to be linear, meaning that they are independent of the level of determinants. We find interesting nonlinearities in the determinants of child mortality, which are important for formulating policies to reduce child mortality.

Further research will address the analysis of structural differences in nonlinearities in the relationship between child mortality and socioeconomic characteristics between Sub-Saharan African and other developing regions.

A Appendix

A.1 Analyzed countries

Country	ISO-Code	Survey-Year	Observations
Benin	204	2006 (5)	13,665
		2001 (4)	4,823
		1996 (3)	2,736
Burkina Faso	854	2003 (4)	10,067
		1998 (3)	4,370
		1992 (2)	4,299
Cameroon	120	2004 (4)	3,303
		1998 (3)	1,872
Chad	148	2004 (4)	4,257
		1996 (3)	5,280
Comoros	174	1996 (3)	954
Ethiopia	231	2005 (5)	5,031
		2000 (4)	10,814
Ghana	288	2008 (5)	3,121
		2003 (4)	3,666
		1998 (4)	2,217
		1993 (3)	1,862
Guinea	324	2005 (5)	3,528
		1999 (4)	4,241
Kenya	404	2009 (5)	5,214
		2003 (4)	4,321
		1998 (3)	3,405
		1993 (3)	4,398
Lesotho	426	2009 (6)	1,506
		2004 (4)	1,478
Liberia	430	2007 (5)	4,885
Madagascar	450	2004 (4)	4,993
		1997 (3)	2,993
Malawi	454	2004 (4)	8,347
		2000 (4)	9,557
		1992 (2)	3,068

Table A.1: List of countries with survey years and number of observations. The number in paranthesis indicates the respective survey wave.

Country	**ISO-Code**	**Survey-Year**	**Observations**
Mali	466	2006 (5)	13,099
		2001 (4)	11,750
		1995 (3)	6,154
Mozambique	508	2003 (4)	8,352
		1997 (3)	3,001
Namibia	516	2007 (5)	4,150
		1992 (2)	1,996
Niger	562	2006 (5)	4,210
		1998 (3)	4,598
		1992 (2)	4,859
Nigeria	566	2008 (5)	24,604
		2003 (4)	5,117
		1999 (4)	2,221
Rwanda	646	2005 (5)	3,766
		2000 (4)	6,184
Senegal	686	2005 (4)	2,611
		1992 (2)	3,862
Tanzania	834	2004 (4)	6,928
		1996 (3)	4,276
		1992 (2)	5,140
Togo	768	1998 (3)	4,029
Uganda	800	2006 (5)	2,472
		2000 (4)	5,200
		1995 (3)	3,811
Zambia	894	2007 (5)	5,240
		2001 (4)	5,765
		1996 (3)	4,602
		1992 (2)	4,114
Zimbabwe	716	2006 (5)	4,397
		1999 (4)	3,039
		1994 (3)	1,903

Table A.1 (cont.): List of countries with survey years and number of observations. The number in paranthesis indicates the respective survey wave.

References

Adebayo, Samson B. and Ludwig Fahrmeir, "Analyzing Child Mortality in Nigeria with Geoadditive Discrete Time Survival Methods", *Statistics in Medicine* 24 (2005), 709–728.

Adeyemi, S L, U A Raheem, and F B Olorunfemi, "Under-Five Mortality and the Environment of Health in the Third World: A Nigerian Example", *Journal of Human Ecology* 24 (2008), 117–124.

Ainsworth, Martha and Deon Filmer, "Inequalities in children's schooling: AIDS, orphanhood, poverty, and gender", *World Development* 34 (2006), 1099–1128.

Almond, Douglas and Janet Currie, "Human capital development before age five", Working Paper 15827, National Bureau of Economic Research (2010).

Armstrong-Schellenberg, Joanna R. M., Rose Nathan, Salim Abdulla, Oscar Mukasa, Tanya J. Marchant, Marcel Tanner, and Christian Lengeler, "Risk factors for child mortality in rural Tanzania", *Tropical Medicine and International Health* 7 (2002), 506–511.

Balk, Deborah, Tom Pullum, Adam Storeygard, Fern Greenwell, and Melissa Neuman, "Spatial Analysis of Childhood Mortality in West Africa", DHS Geographic Studies 1, ORC Macro and Center for International Earth Science Information Network (CIESIN) (2003).

Becher, Heiko, Olaf Müller, Albrecht Jahn, Adjima Gbangou, Gisela Kynast-Wolf, and Bocar Kouyaté, "Risk factors of infant and child mortality in rural Burkina Faso", *Bulletin of the World Health Organization* 82 (2004), 265–273.

Bhalotra, Sonia and Samantha B. Rawlings, "Intergenerational persistence in health in developing countries: The penalty of gender inequality?", *Journal of Public Economics* 95 (2011), 286–299.

Bicego, George, Shea Rutstein, and Kiersten Johnson, “Dimensions of the emerging orphan crisis in sub-Saharan Africa”, *Social Science & Medicine* 56 (2003), 1235–1247.

Bicego, George T. and J. Ties Boerma, “Maternal education and child survival: A comparative study of survey data from 17 countries”, *Social Science & Medicine* 36 (1993), 1207–1227.

Blunch, Niels-Hugo, “Staying Alive: Adult Literacy Programs and Child Mortality in Rural Ghana”, *World Development* 42 (2013), 114–126.

Bollen, Kenneth A., Jennifer L. Glanville, and Guy Stecklov, “Economic status proxies in studies of fertility in developing countries: does the measure matter?”, *Population Studies* 56 (2002), 81–96.

Caldwell, John C., “Education as a Factor in Mortality Decline: An Examination of Nigerian Data”, *Population Studies* 33 (1979), 395–413.

Caldwell, John C., “Mass education as a determinant of mortality decline”, in John C. Caldwell and Gigi Santow, eds., *Selected readings in the cultural, social and behavioural determinants of health* (The Australian National University. Canberra, 1989), pp. 101–109.

Caldwell, John C. and P. McDonald, “Influence of maternal education on infant and child mortality: levels and causes”, *Health Policy and Education* 2 (1982), 251–267.

Caldwell, John C., P.H. Reddy, and Pat Caldwell, “The Social Component of Mortality Decline: An Investigation of South India Employing Alternative Methodologies”, *Population Studies* 37 (1983), 185–205.

Case, Anne, Angela Fertig, and Christina Paxson, “The lasting impact of childhood health and circumstance”, *Journal of Health Economics* 24 (2005), 365–389.

Currie, Janet, "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development", *Journal of Economic Literature* 47 (2009), 87–122.

de Onis, Mercedes and Jean-Pierre Habicht, "Anthropometric reference data for international use: recommendations from a World Health Organization Expert Committee", *American Journal of Clinical Nutrition* 64 (1996), 650–658.

de Onis, Mercedes and Ray Yip, "The WHO growth chart: historical considerations and current scientific issues", *Bibliotheca Nutritio et Dieta* 53 (1996), 74–89.

Engle, Patrice L., Purnima Menon, and Lawrence Haddad, "Care and Nutrition: Concepts and Measurement", *World Development* 27 (1999), 1309–1337.

Fahrmeir, Ludwig, Thomas Kneib, Stefan Lang, and Brian Marx, *Regression: Models, Methods and Applications* (Berlin Heidelberg: Springer-Verlag, 2013).

Fay, Marianne, Danny Leipziger, Quentin Wodon, and Tito Yepes, "Achieving Child-Health-Related Millenium Development Goals: The Role of Infrastructure", *World Development* 33 (2005), 1267–1284.

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

Gemperli, A, P Vounatsou, I Kleinschmidt, M Bagayoko, C Lengeler, and T Smith, "Spatial patterns of infant mortality in Mali: the effect of malaria endemicity", *American Journal of Epidemiology* 159 (2004), 64–72.

Hanf, Matthieu, Mathieu Nacher, Chantal Guihenneuc, Pascale Tubert-Bitter, and Michel Chavance, "Global determinants of mortality in under 5s: 10 year worldwide longitudinal study", *BMJ* 347 (2013), f6427.

Harttgen, Kenneth, Stephan Klasen, and Sebastian Vollmer, “An African Growth Miracle? Or: What Do Asset Indices Tell Us about Trends in Economic Performance”, *Review of Income and Wealth* 59 (2013a), S37–S61.

Harttgen, Kenneth, Stephan Klasen, and Sebastian Vollmer, “Economic Growth and Child Undernutrition in Africa”, *Population and Development Review* 39 (2013b), 397–412.

Harttgen, Kenneth and Mark Misselhorn, “A Multilevel Approach to Explain Child Mortality and Undernutrition in South Asia and Sub-Saharan Africa”, Ibero America Institute for Economic Research (IAI) Discussion Papers 152, Ibero-America Institute for Economic Research (2006).

Heston, Alan, Robert Summers, and Bettina Aten, “Penn World Table Version 7.1”, (2012). Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hobcraft, J.N., J.W. McDonald, and S.O. Rutstein, “Socio-economic Factors in Infant and Child Mortality: A Cross-national Comparison”, *Population Studies* 38 (1984), 193–223.

Jain, Anil K., M N Murty, and Patrick J. Flynn, “Data Clustering: A Review”, *ACM Computing Surveys* 31 (1999), 264–323.

Kandala, Ngianga B., Ludwig Fahrmeir, Stephan Klasen, and Jan Priebe, “Geo-Additive Models of Childhood Undernutrition in Three Sub-Saharan African Countries”, *Population, Space and Place* 15 (2009), 461–473.

Kandala, Ngianga B. and Gebrenegus Ghilagaber, “A Geo-Additive Bayesian Discrete-Time Survival Model and its Application to Spatial Analysis of Childhood Mortality in Malawi”, *Quality and Quantity* 40 (2006), 935–957.

Kazembe, Lawrence, Aileen Clarke, and Ngianga-Bakwin Kandala, “Childhood mortality in sub-Saharan Africa: cross-sectional insight into small-scale geographical inequalities from Census data”, *BMJ Open* 2 (2012), e001421.

Kembo, Joshua and Jeroen K Van Ginneken, “Determinants of infant and child mortality in Zimbabwe: Results of multivariate hazard analysis”, *Demographic Research* 21 (2009), 367–384.

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

Krivobokova, Tatyana, Thomas Kneib, and Gerda Claeskens, “Simultaneous confidence bands for penalized spline estimators”, *Journal of the American Statistical Association* 105 (2010), 852–863.

Kyei, Kwabena A., “Determinants of Childhood Mortality in South Africa: Using Categorical Data Modelling”, *Journal of Human Ecology* 37 (2012), 47–56.

Lang, Stefan, Nikolaus Umlauf, Peter Wechselberger, Kenneth Harttgen, and Thomas Kneib, “Multilevel Structured Additive Regression”, *Statistics and Computing* 24 (2014), 223–238.

Majumder, Abul Kashem, Marian May, and Prakash Dev Pant, “Infant and Child Mortality Determinants in Bangladesh: Are they Changing?”, *Journal of Biosocial Science* 29 (1997), 385–399.

Manda, Samuel O.M., “Birth intervals, breastfeeding and determinants of childhood mortality in Malawi”, *Social Science & Medicine* 48 (1999), 301–312.

Mesike, C G and J N Mojekwu, “Environmental Determinants of Child Mortality in Nigeria”, *Journal of Sustainable Development* 5 (2012), 65–75.

Mosley, Wiley Henry and Lincoln C. Chen, “An Analytical Framework for the Study of Child Survival in Developing Countries”, *Population and Development Review* 10 (1984), 25–45.

Mutunga, Clive J., “Environmental Determinants of Child Mortality in Kenya”, UNO-WIDER Research Paper 2007/83, United Nations University. World Institute for Development Economics Research (2007).

Nuwaha, Fred, Juliet Babirye, Olico Okui, and Natal Ayiga, “Understanding socio-economic determinants of childhood mortality: A retrospective analysis in Uganda”, *BMC Research Notes* 4 (2011), 484.

Ogunjuyigbe, Peter O., “Under-Five Mortality in Nigeria: Perception and Attitudes of the Yorubas Towards the Existence of Abiku”, *Demographic Research* 11 (2004), 43–56.

Ojikutu, Rasheed Kola, “Pattern of Under-Five Deaths in Lagos State, Nigeria”, *Sudanese Journal of Public Health* 3 (2008).

Ojikutu, Rasheed Kola, “Malaria and child survival in Nigeria: beyond the stethoscope”, *International Journal of Academic Research* 2 (2010).

Pelletier, D L and E A Frongillo, “Changes in child survival are strongly associated with changes in malnutrition in developing countries”, *The Journal of Nutrition* 133 (2003), 107–119.

Pritchett, Lant and Lawrence H. Summers, “Wealthier is Healthier”, *The Journal of Human Resources* 31 (1996), 841–868.

Sahn, David E. and David C. Stifel, “Exploring Alternative Measures of Welfare in the Absence of Expenditure Data”, *Review of Income and Wealth* 49 (2003), 463–489.

Sastry, Narayan, “Trends in Socioeconomic Inequalities in Mortality in Developing Countries: The Case of Child Survival in São Paulo, Brazil”, *Demography* 41 (2004), 443–464.

Schellenberg, Joanna A., Cesar G. Victoria, Adiel Mushi, Don de Savigny, David Schellenberg, Hassan Mshinda, and Jennifer Bryce, “Inequities among the very poor: health care for children in rural southern Tanzania”, *The Lancet* 361 (2003), 561–566.

Schultz, T Paul, “Studying the Impact of Household Economic and Community Variables on Child Mortality”, *Population and Development Review* 10 (1984), 215–235.

Smith-Greenaway, Emily, “Maternal Reading Skills and Child Mortality in Nigeria: A Reassessment of Why Education Matters”, *Demography* 50 (2013), 1551–1561.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde, “Bayesian Measures of Model Complexity and Fit”, *Journal of the Royal Statistical Society, Series B* 64 (2002), 583–639.

Stifel, David and Luc Christiaensen, “Tracking poverty over time in the absence of comparable consumption data”, *World Bank Economic Review* 21 (2007), 317–341.

Sullivan, Jeremiah M., Shea Oscar Rutstein, and George T. Bicego, “Infant and Child Mortality”, DHS Comparative Studies 15, Macro International Inc. (1994).

Tarozzi, Alessandro and Aprajit Mahajan, “Child Nutrition in India in the Nineties: a Story of Increased Gender Inequality?”, Working Papers 05-06, Duke University, Department of Economics (2005).

Umlauf, Nikolaus, Daniel Adler, Thomas Kneib, Stefan Lang, and Achim Zeileis, “Structured Additive Regression Models: An R Interface to BayesX”, Working Papers 2012-10, Faculty of Economics and Statistics, University of Innsbruck (2013a). Revised for *Journal of Statistical Software*.

Umlauf, Nikolaus, Nadja Klein, Achim Zeileis, and Stefan Lang, *BayesR: Bayesian Regression in R* (2013b). R package version 0.1-1.

UNICEF, “Strategy for improved nutrition of children and women in developing countries”, (1990). New York.

UNICEF, “Levels & Trends in Child Mortality”, Report 2013, United Nations Children’s Fund (2013).

van de Poel, Ellen, Owen O'Donnell, and Eddy van Doorslaer, "What Explains the Rural-Urban Gap in Infant Mortality: Household or Community Characteristics?", *Demography* 46 (2009), 827–850.

Wang, L, "Environmental Determinants of Child Mortality: Empirical Results from the 2000 Ethiopia DHS", (2003). World Bank. Washington DC.

WHO, "Physical Status: The Use and Interpretation of Anthropometry", WHO Technical Report Series 854, World Health Organization (1995).

WHO, "Obesity: Preventing and Managing the Global Epidemic", WHO Technical Report Series 894, World Health Organization (2000).

WHO, "WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age", URL `http://www.who.int/childgrowth/standards/technical_report/en/` (2006). Geneva.

Wolpin, Kenneth, "Determinants and consequences of the mortality and health of infants and children", in Mark Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, vol. 1A, chap. 10 (Elsevier, 1997). Amsterdam, North Holland.

World Bank, "World Development Report 2007: Development and the Next Generation", Working papers, The World Bank (2006). Oxford University Press, Washington, DC.

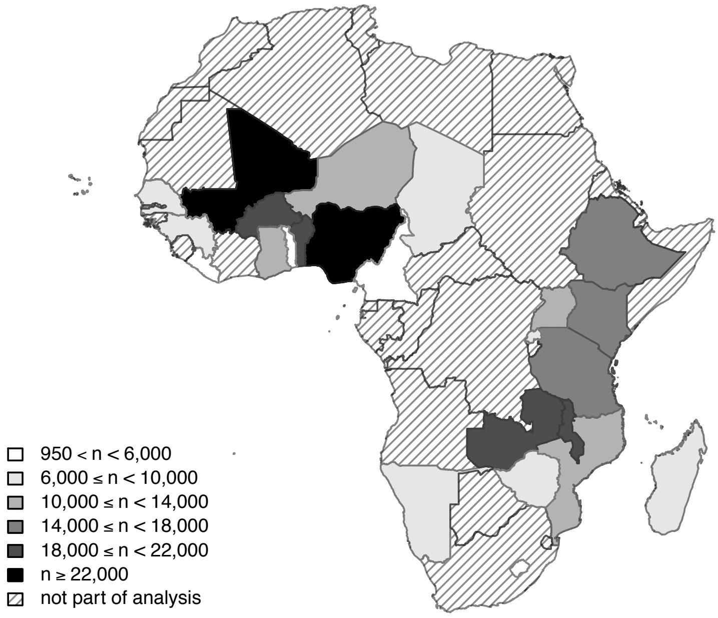

Figure 1: Map of Africa indicating the number of observations in each country.

Variable	Definition	%, Mean (Sd)		Level
dead5	1 = Child is dead	18 %		level-1
	0 = Child is alive	82 %		
adults	Number of adults	3		level-1
agemother	Current age of the mother at the time of interview	32.8	(6.69)	level-1
mbmi	Body mass index of the mother	22.02	(3.35)	level-1
ai_drm	Asset index (deviation from regional mean)	-0.18	(0.78)	level-1
gender	Sex of child			level-1
	1 = Child is male	50.7 %		
	-1 = Child is female	49.3 %		
fhh	Female headed household			level-1
	1 = yes	47.01 %		
	-1 = no (male headed)	52.99 %		
firstbchild	Child is first born			level-1
	1 = yes	21.3 %		
	-1 = no	78.7 %		
deadsibling	Child has a dead sibling			level-1
	1 = yes	16.66 %		
	-1 = no	83.34 %		
mliterate	Mother is literate			level-1
	1 = yes	18.46 %		
	-1 = no	81.54 %		
urban	Place of residence			level-1
	1 = urban	23.18 %		
	-1 = rural	76.82 %		
wavet	Survey wave t			level-1
	wave2: 1990-1992	8.66 %		
	wave3: 1993-1998	19.08 %		
	wave4: 1998-2005	40.94 %		
	wave5: 2005-2009	30.84 %		
	wave6: 2009-2011	0.48 %		
regid	Identification code for regions (combination of country code and number of regions)			level-2
stunting_rm	Z-score stunting (regional mean)			level-2
country	ISO-code (classification *3166-1 numeric*)			level-3
rgdp	Real gross domestic product per capita at 2005 constant prices (source: PWT 7.0 2012)			level-3

Table 1: List of variables. Source: Demographic and Health Surveys, Penn World Tables; calculation by the authors.

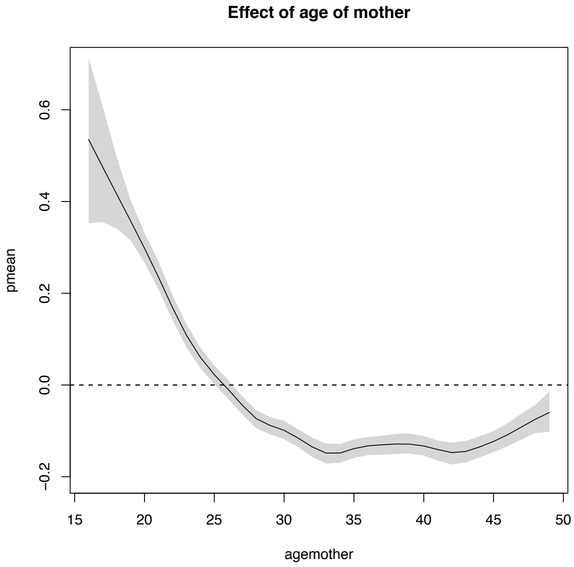

Figure 2: Effect plot of $f_1(agemother)$.

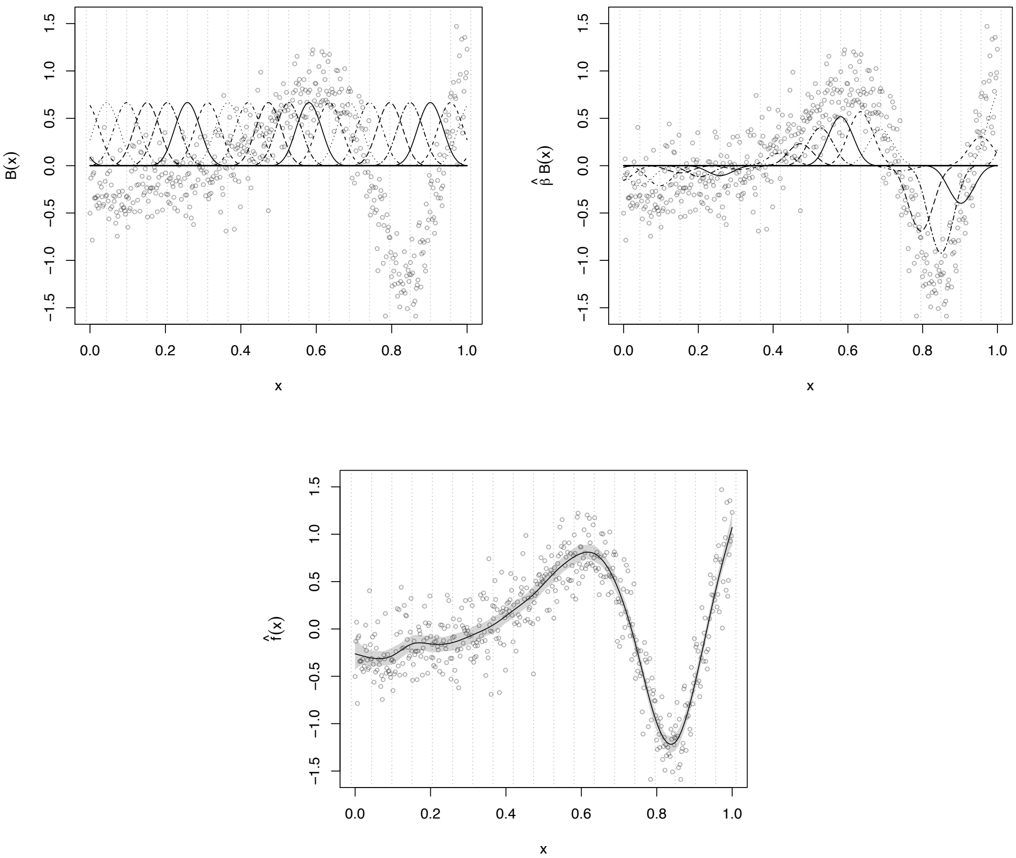

Figure 3: Illustration of the approximation of a nonlinear relationship using Penalized splines.

Variable	Description	Type	Expected effect
gender	Sex of child	discrete	+
fhh	Female headed household	discrete	o
firstbchild	Child is first born	discrete	+
deadsibling	Child has a dead sibling	discrete	+
mliterate	Mother is literate	discrete	−
urban	Place of residence	discrete	−
wave2	Survey wave 1990 - 1992	discrete	+
wave3	Survey wave 1993 - 1998	discrete	+
wave5	Survey wave 2005 - 2009	discrete	−
wave6	Survey wave 2009 - 2011	discrete	−
agemother	Current age of the mother at the time of interview	continuous	U-shaped
adults	Number of adults in household	continuous	monotonically increasing
mbmi	Body Mass Index of Mother	continuous	U-shaped
ai_drm	Asset Index (deviation from regional average)	continuous	monotonically decreasing
stunting_rm	Z-score stunting (regional average)	continuous	monotonically decreasing
rgdp	Real gross domestic product per capita	continuous	slightly decreasing

Table 2: Expected effects of the explanatory variables on child mortality. The effects are expected to be: positive (+), negative (−) or no strong assumptions (o) in case of discrete covariates. Regarding the nonlinearly modelled continuous covariates we expect: U-shaped, monotonically increasing or monotonically decreasing relationships.

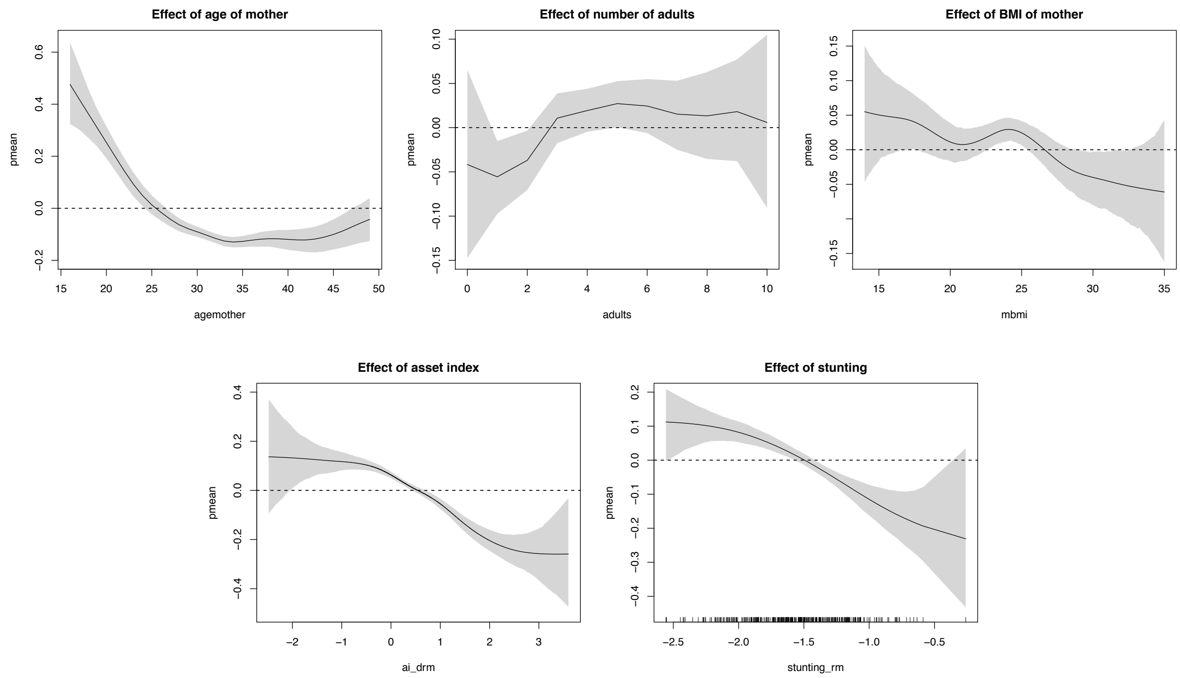

Figure 4: Effects of age of mother, number of adults in household, BMI of mother and asset index at level-1 and stunting at level-2. The covariate at level-2 represents the regional mean of the Z-score (stunting). The shaded area depicts the 95% pointwise credible intervals.

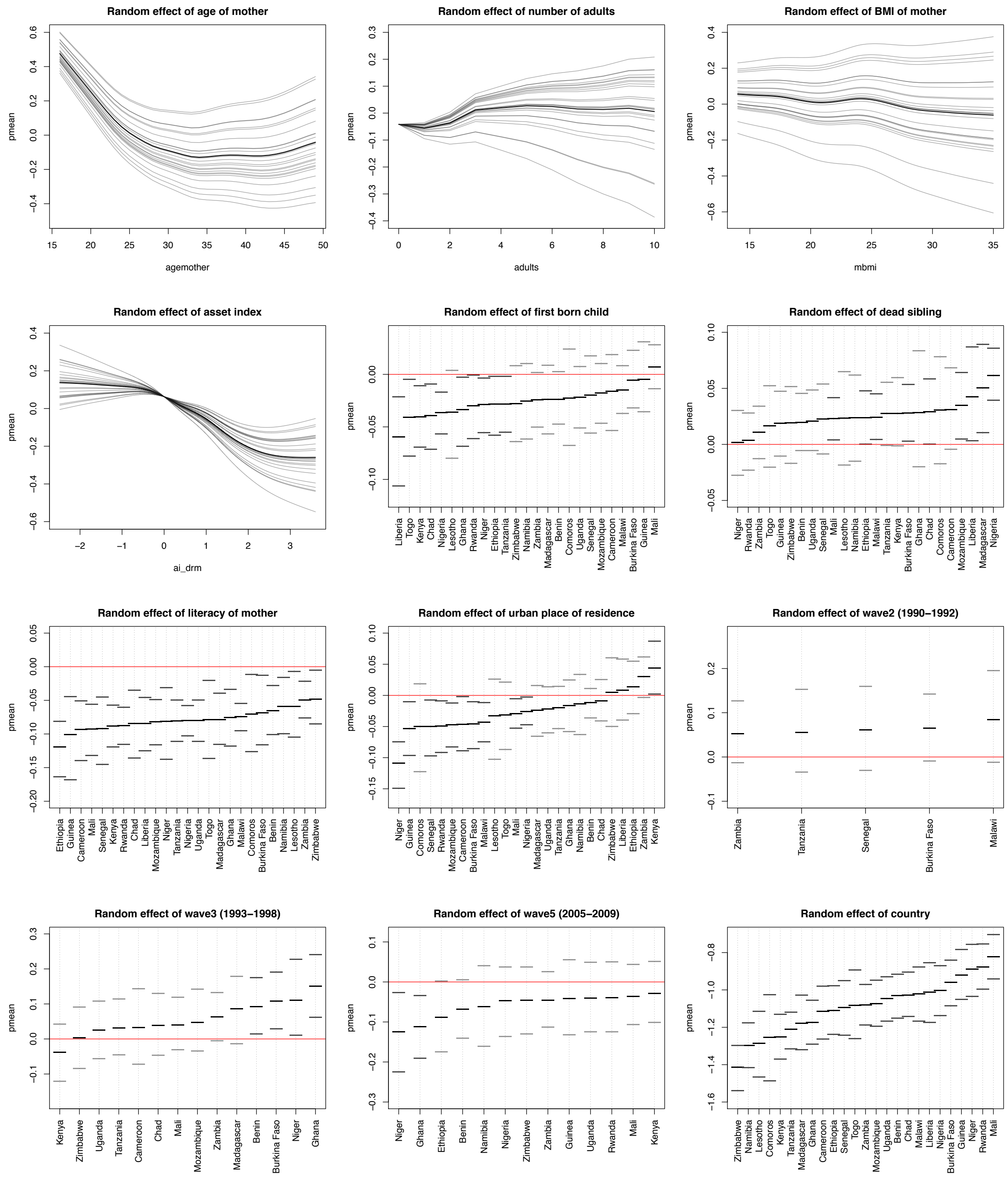

Figure 5: Country random effects of continuous and categorical covariates. In case of continuous covariates we depict the main effect (black line) as well as the 25 country-specific regression curves (grey lines). In case of categorical covariates, we plot the posterior means as black bars in ascending order together with the 95% credible intervals defined by the lower and upper grey colored bars. The effects are assumed to be significant if the credible interval does not contain zero. We indicate significance by dark grey colored, insignificance by light grey colored bars.

Notes

[1]Before the new growth standard (WHO, 2006) anthropometric indicators were based on the Health Statistics/World Health Organization (NCHS/WHO) growth reference that had been recommended for international use since the late 1970s (WHO, 1995). The data to construct the reference standard came from a longitudinal study of children of European ancestry from a single community in the USA where children were measured every three months. The limitations of the NCHS/WHO reference standard to adequately represent early childhood growth led to the construction of a new early childhood growth standard (de Onis and Yip, 1996; de Onis and Habicht, 1996). The data used to construct the new WHO (2006) growth standard comes from six countries: Brazil, Ghana, India, Norway, Oman and the USA, which makes the standard applicable for international comparisons. Within these countries the healthy population is selected for the study to minimize the impact of environmental variations.

[2]Similarly, underweight is defined as a Z-score less than minus two for weight relative to children of the same sex and age in the reference population. Wasting is defined as a Z-score less than minus two for weight-to-height relative to children of the same sex and age in the reference population. Biologically impossible values are defined by the WHO for height (stunting) as Z-scores $<$-6 or $>$6; for weight (underweight) as $<$-6 or $>$5; and for weight for height (wasting) as $<$-5 or $>$5.

[3]For example, there is a large body of literature that uses an asset index to explain inequalities in child malnutrition (e.g Sahn and Stifel, 2003; Tarozzi and Mahajan, 2005), child mortality (e.g. Sastry, 2004) when data on income or expenditure is not available, or educational outcomes (e.g. Ainsworth and Filmer, 2006; Bicego et al., 2003), health outcomes (e.g. Bollen et al., 2002; Schellenberg et al., 2003). Further, asset indices are used to analyze changes and determinants of poverty (Harttgen et al., 2013a,b; Stifel and Christiaensen, 2007; World Bank, 2006).

[4]The Penn World Tables provide data for GDP per capita that is comparable across countries and time as this measure of income per capita accounts for Purchasing Power Parity using 2005 International Comparison Program prices as the benchmark and adjustments for exchange rate variation.

[5]DHS wave 1 is from 1985 to 1990, DHS wave 2 is from 1990 to 1992, DHS wave 3 is from 1993 to 1998, DHS wave 4 is from 1998 to 2005, DHS wave 5 is from 2005 to 2009, and DHS wave 6 is from 2009 to 2011. Since no information on anthropometric indicators is available before 1990, DHS wave 1 is dropped from the sample.

[6]More formally, f forms a polynomial spline of degree three, which is a polynomial in each interval and two times continuously differentiable at the interval boundaries or knots.

[7]For details see Table A.1 in the Appendix.

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

http://eeecon.uibk.ac.at/wopec/

2015-03 **Kenneth Harttgen, Stefan Lang, Judith Santer:** Multilevel modelling of child mortality in Africa

2015-02 **Helene Roth, Stefan Lang, Helga Wagner:** Random intercept selection in structured additive regression models

2015-01 **Alice Sanwald, Engelbert Theurl:** Out-of-pocket expenditures for pharmaceuticals: Lessons from the Austrian household budget survey

2014-32 **Esther Blanco, Tobias Haller, James M. Walker:** Externalities in appropriation: Responses to probabilistic losses

2014-31 **Michael Kirchler, Stefan Palan:** Friendliness pays off! Monetary and immaterial gifts in consumer-salesperson interactions

2014-30 **Alice Sanwald, Engelbert Theurl:** Out-of-pocket expenditures of private households for dental services - Empirical evidence from Austria

2014-29 **Stefan Borsky, Esther Blanco:** Setting one voluntary standard in a heterogeneous Europe - EMAS, corruption and stringency of environmental regulations

2014-28 **Eberhard Feess, Christian Grund, Markus Walzl, Ansgar Wohlschlegel:** Competing trade mechanisms and monotone mechanism choice

2014-27 **Esther Blanco, E. Glenn Dutcher, Tobias Haller:** To mitigate or to adapt? Collective action under asymmetries in vulnerability to losses

2014-26 **Alice Sanwald, Thomas Schober:** Follow your heart: Survival chances and costs after heart attacks - An instrumental variable approach

2014-25 **Eric Mayer, Sebastian Rüth, Johann Scharler:** Total factor productivity and the propagation of shocks; Empirical evidence and implications for the business cycle

2014-24 **Susanne Berger, Herbert Stocker, Achim Zeileis:** Innovation and institutional ownership revisited: An empirical investigation with count data models

2014-23 **James C. Cox, Rudolf Kerschbamer, Daniel Neururer:** What is trustworthiness and what drives it?

2014-22 **Michael Kirchler, Caroline Bonn, Jürgen Huber, Michael Razen:** The "Inflow-Effect" - Trader inflow and bubble formation in asset markets

2014-21 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Donations, risk attitudes and time preferences: A study on altruism in primary school children *forthcoming in* *Journal of Economic Behavior and Organization*

2014-20 **Christian Kleiber, Achim Zeileis:** Visualizing count data regressions using rootograms

2014-19 **Matthias Siller, Christoph Hauser, Janette Walde, Gottfried Tappeiner:** The multiple facets of regional innovation

2014-18 **Carmen Arguedas, Esther Blanco:** On fraud and certification of corporate social responsibility

2014-17 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Home victory for Brazil in the 2014 FIFA World Cup

2014-16 **Andreas Exenberger, Andreas Pondorfer, Maik H. Wolters:** Estimating the impact of climate change on agricultural production: accounting for technology heterogeneity across countries

2014-15 **Alice Sanwald, Engelbert Theurl:** Atypical employment and health: A meta-analysis

2014-14 **Gary Charness, Francesco Feri, Miguel A. Meléndez-Jiménez, Matthias Sutter:** Experimental games on networks: Underpinnings of behavior and equilibrium selection *slightly revised version forthcoming in* *Econometrica*

2014-13 **Uwe Dulleck, Rudolf Kerschbamer, Alexander Konovalov:** Too much or too little? Price-discrimination in a market for credence goods

2014-12 **Alexander Razen, Wolgang Brunauer, Nadja Klein, Thomas Kneib, Stefan Lang, Nikolaus Umlauf:** Statistical risk analysis for real estate collateral valuation using Bayesian distributional and quantile regression

2014-11 **Dennis Dlugosch, Kristian Horn, Mei Wang:** Behavioral determinants of home bias - theory and experiment

2014-10 **Torsten Hothorn, Achim Zeileis:** partykit: A modular toolkit for recursive partytioning in R

2014-09 **Rudi Stracke, Wolfgang Höchtl, Rudolf Kerschbamer, Uwe Sunde:** Incentives and selection in promotion contests: Is it possible to kill two birds with one stone? *forthcoming in* *Managerial and Decision Economics*

2014-08 **Rudi Stracke, Wolfgang Höchtl, Rudolf Kerschbamer, Uwe Sunde:** Optimal prizes in dynamic elimination contests: Theory and experimental evidence *forthcoming in Journal of Economic Behavior and Organization*

2014-07 **Nikolaos Antonakakis, Max Breitenlechner, Johann Scharler:** How strongly are business cycles and financial cycles linked in the G7 countries?

2014-06 **Burkhard Raunig, Johann Scharler, Friedrich Sindermann:** Do banks lend less in uncertain times?

2014-05 **Julia Auckenthaler, Alexander Kupfer, Rupert Sendlhofer:** The impact of liquidity on inflation-linked bonds: A hypothetical indexed bonds approach

2014-04 **Alice Sanwald, Engelbert Theurl:** What drives out-of pocket health expenditures of private households? - Empirical evidence from the Austrian household budget survey

2014-03 **Tanja Hörtnagl, Rudolf Kerschbamer:** How the value of information shapes the value of commitment or: Why the value of commitment does not vanish

2014-02 **Adrian Beck, Rudolf Kerschbamer, Jianying Qiu, Matthias Sutter:** Car mechanics in the lab - Investigating the behavior of real experts on experimental markets for credence goods *forthcoming in Journal of Economic Behavior and Organization*

2014-01 **Loukas Balafoutas, Adrian Beck, Rudolf Kerschbamer, Matthias Sutter:** The hidden costs of tax evasion - Collaborative tax evasion in markets for expert services

University of Innsbruck

Working Papers in Economics and Statistics

2015-03

Kenneth Harttgen, Stefan Lang, Judith Santer

Multilevel modelling of child mortality in Africa

Abstract
Whereas child mortality has been decreased globally in the last 20 years, high levels persist in Sub-Saharan Africa. This paper analyzes child mortality in 25 Sub-Saharan countries based on household survey data. We employ a new multilevel approach with structured additive predictor within the Bayesian framework. This allows us to take into account the hierarchical data structure and use the heterogeneity within and between countries as well as to assess non-linearities in the relationship between child mortality and socio-economic determinants. We find that household's economic well-being, mother's education and age, and geographical regions strongly influence child mortality risks.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)