

Harttgen, Kenneth; Lang, Stefan; Santer, Judith; Seiler, Johannes

Working Paper

Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa

Working Papers in Economics and Statistics, No. 2017-15

Provided in Cooperation with:

Institute of Public Finance, University of Innsbruck

Suggested Citation: Harttgen, Kenneth; Lang, Stefan; Santer, Judith; Seiler, Johannes (2017) : Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa, Working Papers in Economics and Statistics, No. 2017-15, University of Innsbruck, Research Platform Empirical and Experimental Economics (eeecon), Innsbruck

This Version is available at:

<https://hdl.handle.net/10419/180166>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa

Kenneth Harttgen, Stefan Lang, Judith Santer,
Johannes Seiler

Working Papers in Economics and Statistics

2017-15

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.

Modeling Under-5 Mortality through Multilevel Structured Additive Regression with Varying Coefficients for Asia and Sub-Saharan Africa

Kenneth Harttgen[†], Stefan Lang[‡], Judith Santer[‡], and Johannes Seiler^{*‡}

[†]ETH Zurich

[‡]University of Innsbruck

August 22, 2017

Despite improvements in global child health within the last three decades, under-5 mortality remains significantly high in Sub-Saharan and Asia. Both regions did not achieve the MDG target of reducing under-5 mortality by two thirds by 2015. The underlying causes of under-5 mortality differ significantly between countries and between regions, which highlights the need to expand our understanding of the determinants of child health in developing countries. By comparing the two geographic regions of the world with the highest under-5 mortality rates, we aim to gain new insights, and bring out potential differences between these regions and the causes of under-5 mortality. In addition, we aim to identify non-linear relationships between under-5 mortality and specific explanatory variables. We analyze a large data set consisting of 35 Sub-Saharan-African countries, and 13 Asian countries, using a multilevel discrete time survival model that takes advantage of a recently developed multilevel framework with structured additive predictor in a Bayesian setting. We analyze data from 131 individual surveys from 1992 to 2015, allowing for potential non-linear effects and cluster specific heterogeneity within models. We find strong non-linear effects for the baseline hazard, the household size, the age of the mother, the BMI of the mother, and the birth order of the child. Additionally, we find considerable differences in determinants between Asian and Sub-Saharan Asian countries.

Keywords: Child mortality · Asia · Sub-Sahara Africa · multilevel STAR models · Bayesian inference

JEL-Classification: C11, I15, O57

*J. S. acknowledges financial support from a doctoral scholarship by the University of Innsbruck.

1. Introduction

Global child health has been improved throughout the last three decades in low and middle-income countries. Most regions saw under-5 mortality rates decline significantly since 1990. This trend has been confirmed by the UNIGME, who testified that substantial global progress has been observed in decreasing child mortality since 1990 (You et al., 2012, 2014; You, New and Wardlaw, 2015). Most regions experience an accelerated decline in under-5 mortality since 2000 compared to the period 1990-2000. The annual rate of reduction increased in Sub-Saharan Africa from 1.6% in 1990-2000 to 4.1% in 2000-2015, in Central Asia from 1.4% to 4.6%, and in Southern Asia from 3.2% to 3.6%. However, despite this significant progress towards achieving MDG 4, high under-5 mortality rates are still concentrated in these regions, where rates remain higher as targeted by the MDG for 2015 (You, New and Wardlaw, 2015).

Therefore, analyzing and understanding the link between the health outcomes of children, and the determinants of child mortality is of great interest for future policies and aid. A better understanding of this relationship, and applying this knowledge to future policy can effectively help to achieve the newly formulated SDG¹ in the long run. In addition, through more purposeful aid the health outcome of children can be improved in the short-term.

The empirical literature on studying determinants of child health in general, and under-5 mortality in particular has identified several factors of child wellbeing: child health is influenced by individual and household characteristics (e.g. residency, structure and composition of the household, education, wealth, nutrition, immunization, hygiene, care) and environmental and community or regional conditions (e.g. infrastructure, climate, disease environment) that directly and indirectly determine proximate or intimate input variables for child health outcomes (Mosley and Chen, 1984; Schultz, 1984; Wolpin, 1997).

Based on this theoretical considerations, a broad stream of empirical literature exists, which uses regression methods to analyze survival outcomes of infants, and children's under a certain age (e.g. 36 month, 60 month) in developing countries, as it can be seen as reverse indicator of child health (see for instance Boco, 2010; Hobcraft et al., 1985; Houweling et al., 2005; Kandala and Ghilagaber, 2006; Liu et al., 2016; Misselhorn and Harttgen, 2006; Rutstein, 2000, 2005; Wang, 2003; You et al., 2010). Therefore, our main goal is to systematically analyze under-5 mortality, comparing determinants across geopolitical regions, and allowing for non-linear covariate effects using a multilevel regression model.

Building on this literature, in our research, we will allow for possible non-linear covariate effects on under-5 mortality in our analysis, and we will jointly analyze the two geopolitical regions where under-5 was found to be highest. The aim of this study is to better understand why under-5 mortality decreased in such great variation between low and middle-income countries in Sub-Sahara Africa and Asia, and why high rates of

¹The SDG where adopted by the United Nations in 2015 and can be seen as successor of the MDG. In contrast to the MDG they cover not only development related topics, instead the focus of the SDG is on sustainable development

mortality remain in many countries. The underlying causes of under-5 mortality differ significantly between countries and between regions, which highlights the need to expand our understanding of the determinants of child health in developing countries (Black et al., 2003; Jamison et al., 2016). By comparing the two geographic regions of the world with the highest under-5 mortality rates, we aim to gain new insights, and bring out potential differences between these regions, and the causes of under-5 mortality. In addition, we aim to identify non-linear relationships between under-5 mortality and specific explanatory variables. In doing so, we contribute to this strand of literature by aiming to close four shortcomings in the existing literature.

First, we contribute to the literature by simultaneously taking into account possible non-linearities in determinants of under-5 mortality, and heterogeneities across countries. A shortcoming of most existing studies is that they only present linear estimates of determinants of child health. But the mechanisms underlying health transmission suggest that it may be important to allow for non-linearities, which we address with this study. Harttgen et al. (2015) found strong non-linearities in the relationship between under-5 mortality and its determinants. In this article under-5 mortality, and the underlying causes were analyzed for a sample of 25 Sub-Saharan African. Besides heterogeneity across countries, and regions, non-linear effects were found for the age of the mother at birth, the BMI of the mother, and the household's wealth. However, this study did not account for censoring of the data, as a regression model with binary response was estimated which only allowed to include children fully exposed to the risk of dying between birth and the 5th birthday. By using instead a (multilevel) discrete time survival model, it is possible to control for right censoring of the data, model cluster specific heterogeneity, and by including observations from Asia and Sub-Saharan Africa determine differences in the underlying causes by comparing the two geopolitical regions affects most.

Several country studies found non-linearities in the relationship between under-5 mortality and its determinants. For instance, Ayele and Zewotir (2016) found for an Ethiopian sample besides spatial differences, non-linear covariate effects for the household size, the birth order, and the age of the mother at 1st birth. Black et al. (2013) shows that the effect of the nutritional status of a mother is associated with a higher mortality risk of the child. Additionally, they found that the same relation holds when the mother is obese. This relation can be described to be U-shaped for a range of the mother's BMI where the risk of deceasing for the child is lowest in a certain range in the middle of the range of the BMI. However, they did not explicitly model this non-linear behavior. A similar U-shaped pattern is assumed for the age of the mother. Children born by young mother's, might due to limited experience in bringing up children, or their own adolescent show higher mortality risk as children born by mother's in their 20th and early 30th. However, this effect also tends to reverse again for older mother's. This pattern is described and measured for instance by Kandala and Ghilagaber (2006); Kandala et al. (2014); Harttgen et al. (2015).

DHS data does neither provide information on household's income, nor on household's consumption expenditures which would directly reflect the economic status of the household the children belongs to. However, by using the information the DHS provide on

a household's durables and assets, an asset index as proposed by Filmer and Pritchett (2001), and Sahn and Stifel (2003) can be constructed using a principal components analysis. This should approximate the long term economic status of the household. The indicator for the household's wealth, the asset index, is assumed to be monotonically decreasing with child mortality. Also for the years of education of the mother, a monotonically decreasing relationship is assumed. This relationship was already empirically established by (Cleland and van Ginneken, 1988), and can be attributed to two major thoughts. First, women with higher education can better assess relevant information which, in the end, will benefit her children's health and overall status. Second, higher education has a monetary payoff, since the chance to find better jobs increases. However, this relationship needs to be considered carefully, since the educational status can also be seen as approximation of the household's socioeconomic status (Desai and Alva, 1998). Providing new insights on potential non-linearities, meaning that under-5 mortality is dependent on the level of its determining factors. This would contribute to a better understanding of how to reduce under-5 mortality not only in Sub-Saharan Africa.

Second, by using a large set of standardized cross-country household surveys to study determinants of under-5 mortality while taking into account the hierarchical data structure of the micro-data. Most existing studies that analyze individual level data focus on single countries, or a small group of countries within a specific geopolitical region (see for instance Adebayo and Fahrmeir, 2005; Ayele et al., 2015; Chamarbagwala, 2010). Limited empirical evidence exists on the association between child health, and household, child, environmental, and socio-economic characteristics using pooled household survey data from several countries and years at the child level, while simultaneously taking into account the clustered data structure of the survey data, and analyzing non-linearity relationships in determining factors of child health. Studies that account for the clustered data structure are for example Boco (2010); Kuate-Defo and Diallo (2002). While the first study examines community level determinants of under-5 mortality for 28 Sub-Saharan African countries using a multilevel survival model, the second study clusters mortality rates across regions and estimates the community level determinants of child mortality for 27 African countries using DHS data.

A comparison across countries and geographic regions was regarded as not helpful by some authors (e.g. Black et al., 2003), due to different causes, and different magnitude of the effects on under-5 mortality. However, the increased availability of standardized household survey data across countries in recent years, as well as recent theoretical advances in hierarchical modelling in a Bayesian framework (Lang et al., 2014; Belitz et al., 2015) allows us to jointly analyze under-5 mortality in Asia and Sub-Saharan Africa at the individual level. Furthermore it is possible, to account for regional or country specific heterogeneity. We address these issues, by analyzing the largest available cross-country data set at the individual (child level) with a sample size of more than 800,000 children from 48 countries between 1992 and 2015.

Third, we explicitly take into account differences in the determinants of under-5 mortality between geopolitical regions within our model. For example, Klasen (2000) and Misselhorn and Harttgen (2006) have shown that mortality rates in African countries are almost twice as high as in Asian countries, which means that a children born in

Sub-Saharan Africa are confronted with a higher mortality risk than a children growing up in Asia. This might be influenced by different factors associated with child health between Asia and Sub-Sahara Africa, as well as differences in the non-linear nature of these effects. Factors associated with children’s health outcomes vary across this geopolitical regions, and also within countries between regions Black et al. (2003). Which makes it necessary to account for this type of heterogeneity. Additionally, due to possible non-linearities, in countries with a better supply / higher level of a certain factor associated with child mortality the health outcome is better due to this non-linearities. We will take advantage of the data structure and analyze under-5 mortality on an individual level, and simultaneously account for region and country specific heterogeneity in a Bayesian framework.

Fourth, we employ a discrete survival model to address the shortcoming that many studies use a binary variable to study under-5 mortality (usually 0 = alive, 1 = dead) and, hence, do not take into account right censoring of the data. The issue that arises here is that sampled observations have to be fully exposed to under-5 mortality, whereas observations that are not fully exposed to under-5 mortality can not be included in the sample, since otherwise a sampling error would be introduced. Due to this, several studies try to overcome this critique by considering a (discrete time) survival model to account for right censoring of those observations not fully exposed to under-5 mortality (see for instance Adebayo and Fahrmeir, 2005; Kandala and Ghilagaber, 2006; Kandala et al., 2014) however, these studies apply (geoadditive) discrete time survival models only to a specific country. By extending an discrete time survival model, and taking advantage of the recently developed multilevel structured additive regression (STAR) approach (Lang et al., 2014) we are able to systematically analyze under-5 mortality at the individual level, while taking into account the hierarchical structure of the data. Furthermore, this approach takes right censoring of the data into account, and allows us to compare the two geopolitical regions where child mortality was found to be highest. Accounting for these issues, we apply a (multilevel) discrete time survival model. In contrast to the afore mentioned caveat, observations not fully exposed to under-5 mortality can be included in this type of model. DHS provide information on the survival status and birth history of the respondent’s children. Which allows us to apply a discrete time survival model in this analysis.

The remainder of the paper is organized as follows: Section 2 gives an overview of the used data set and a description of the covariates. This is followed by the description of the method. Section 3 includes an overview of the estimated model and the methodology used. Section 4 summarizes the results. Finally, in Section 5 we conclude.

2. Data

2.1. Source of the data

We use nationally representative Demographic and Health Surveys (DHS). DHS are collected by Macro International Inc., Calverton, Maryland in cooperation with local administration and is funded by USAID. The data is collected in waves, with currently the 7th wave ongoing, and normally a period of 5 years between the waves. Currently more than 250 standard DHS for over 90 lower middle income countries (LMICs) are available, and the usual sample size ranging between 5,000 and 30,000 households. DHS are standardized across countries, and collect data on topics such as child health, infant and child mortality, education, and maternal health to household’s durables, and quality of the household’s sanitation and dwelling.

Our study contains information on 824,017 children from 48 countries which were sampled in 131 individual surveys over the period from 1992 to 2015. Additionally, the economic indicators are taken from the PWT (Version 9.0) (Feenstra et al., 2015). Our sample consists of 13 Asian² and 35 Sub-Saharan African countries. These countries can further be divided into into 358 distinct regions. Figure 1 depicts information about the included countries, and countries that could not be included.³

[Please insert Figure 1 about here]

2.2. Dependent variable

Analyzing under-5 mortality is equivalent to model the probability of dying between birth and the 5th birthday (Rutstein and Rojas, 2006). Information on under-5 mortality comes from the birth history of the DHS, which offer basic information for deceased children (sex, data of birth, and date of death) and more detailed information for living children such as the children’s health and nutritional status. In the first type of model, the analysis is restricted to observations that are fully exposed to the risk of dying before the 5th birthday and therefore are not censored. The second type of model, survival models this restriction is not necessary. Both types of models assume a binary response, however to apply discrete time survival models, it is necessary to augment the data, for an example see the methodology manual of **BayesX**. We define an indicator of under-5 mortality for child i that is equal to 0, if the child is still alive in interval t and 1 if the child died in interval t :

$$dead_{it} = \begin{cases} 1, & \text{if } t = t_i, \text{ and } dead_i = 1, \\ 0, & \text{otherwise.} \end{cases}$$

²Asian countries include Central-Asian countries as Armenia, Azerbaijan, Jordan, Kazakhstan, Kyrgyzstan, Turkey, Uzbekistan, South-Asian and South-East-Asian countries as Bangladesh, Cambodia, India, Maldives, Nepal and Pakistan and countries from the Arabian Peninsula geographically located in Asia as Jordan, see Deaton (2007)

³Table A.1 subsumes the number of children and mortality rates per country and survey year.

Taking advantage of a discrete time survival model, our sample is not restricted to observations fully exposed to under-5 mortality, and therefore observations are not limited to the restriction to be born at least five years previous the survey, as suggested by Bhallowa (2006). Applying a discrete time survival model takes censoring into account. This means that the sample is not restricted to children fully exposed to under-5 mortality. Instead, children younger than 60 months at the date of the interview are considered. In Table 1 under-5 mortality rates by continent are presented. Particularly noticeable is that the observed mortality rate of African countries is almost twice as high, when compared to the observed mortality rate Asian countries are faced with. A similar pattern was already pointed out by Klasen (2000) and Misselhorn and Harttgen (2006) who also studied differences in the determinant of under-5 mortality between Sub-Saharan Africa and Asia.

[Please insert Table 1 about here]

2.3. Independent variables

The explanatory variables used in the analysis can be grouped into child-specific factors, environmental and socio-economic characteristics, household-specific and demographic factors, and maternal characteristics. These characteristics are either available as discrete covariates, or as continuous covariates. An early survey on determinants of child mortality and candidates for potential covariates can be found in Mosley and Chen (1984). This contribution added a theoretical framework how to analyze the underlying causes of child mortality in developing countries. The relationship between child health outcomes and the underlying risk factors described in this article are found to be influential by several empirical studies, (see for instance Rutstein, 2000; Black et al., 2003; Adebayo and Fahrmeir, 2005; Rutstein and Rojas, 2006; Boco, 2010).

Child-specific factors. Child-specific determinants include a categorical covariate for the sex of the child. The left hand panel of Figure 2 shows the Nelson-Aalen cumulative hazard by gender and continent, illustrating potential gender and geographical differences. To control for this potential gender gap, including this variable seems to be valid. The same plot suggests that mortality is not evenly distributed over the whole age range from 0 to 59 months. This effect will be captured in the baseline effect of the discrete survival model.⁴

In addition to the sex of the children, a categorical indicator whether an older sibling of observation i died is included into our model, as it is found to be positively related with under-5 mortality. The children's nutritional status is also found to be a major cause of child mortality (see for instance Pelletier and Frongillo, 2003; Black et al., 2003). However, since information on the children's nutritional status is only available for non-deceased children this variable enters the model only at the second level. Unfortunately,

⁴In particular, due to restricted memory, and to make an estimation tractable in **BayesX** we divide the duration time scale into semiannual intervals.

this yields to a potential bias, since the nutritional situation is potentially overestimated and one can assume that deceased children have a lower z-score. Despite this, Alderman et al. (2011) found no evidence for this bias. In their simulation study with national representative data for Indian children, no statistical significant evidence was found that selective mortality influences the anthropometric indicator on a large scale. We include the stunting z-score, which is which is the height for age z-score, as an anthropometric indicator at the regional level in our models.⁵ It is used for two reasons: First according to de Onis et al. (2012), prevalence of malnutrition is higher for stunted children compared to the other two anthropometric indicators. Second, more observations at the regional level are available for the anthropometric indicator for height for age (stunting) in contrast to the anthropometric indicators for weight for age (underweight) and weight for height (wasting). Including instead z-wasting or z-underweight would have reduced the sample to less than 358 regions, as in some regions only z-stunting was observed.

Another factor associated with child health is, whether the child received vaccinations against diseases such as measles, polio, tetanus in the first year of living. Wang (2003); Charmorbagwala et al. (2004), and Boco (2010) for instance, find that with increasing immunization status decreases the risk of dying. The later author also points out, that the immunization status is only observed for non-deceased children, does not allow to include the vaccinations status at an individual level. Accordingly, as recommended by Rutstein and Rojas (2006) we calculate the immunization rate for children age 12-23 months at a regional level⁶.

Maternal characteristics. We include the mother’s nutritional status, measured as the body mass index (BMI), which is found to influence the health of the child (Black et al., 2008, 2013). We assume a U-shape relationship, due to the reasoning explained in the introduction which are also supported by Black et al. (2013). The same relationship is assumed for the age of the mother at the birth of the child. Additionally we include, the number of years of education of the mother. This covariate is assumed to have a decreasing effect on the child’s development status. This can be attributed to two major thoughts. First, women with higher education can better assess relevant information which, in the end, will benefit her children’s health and overall status. Second, higher education has a monetary payoff, since the chance to find better jobs increases. Also the spacing between two births was found to influence the survival chances in the first 5 years (Mosley and Chen, 1984; Rutstein, 2000). Including this information however, features a trade-off, since preceding birth intervals can not observed for first born children, and, therefore, we would systematically omit this group introducing potentially a bias. Hence, we will re-estimate the final model including information on the preceding birth interval and will compare whether the results of the other covariates are in line with the results not including this information.

⁵For a more detailed explanation on the calculation of the z-score see Harttgen et al. (2015).

⁶Included vaccines are: BCG (Bacillus Calmette–Guérin against tuberculosis), measles, and 3 doses of DPT (vaccine against diphtheria, pertussis, and tetanus) and polio vaccine each. For calculating the immunization rate only children ageing 12-23 months are included, since immunization should be completed in the first year.

Environmental socio-economic covariates. Several studies found a significant difference in the health status of children when comparing the location of the household, see for instance Black et al. (2003), and more recently Harttgen and Günther (2011). To control for differences between rural (countryside, villages) and urban (cities, towns) areas, an effect coded covariate is included to capture these potential differences.

Other studies have identified a negative relationship between a country's economic wealth, measured as GDP and not only under-5 mortality, but child mortality in general (see for instance Baird et al., 2010; Bhalotra, 2006). Countries with a higher GDP can potentially invest more in the country's health system, food programs, and schooling and education, which can be translated into decreasing mortality rates.

Moreover, we assume that the included covariate effects are not equal across regions and countries. Hence, we cluster the data geographically according to the observation's affiliation to a particular region and country (Jain et al., 1999; Korenromp et al., 2004). Clustering the observations into regions and countries is particularly useful to model potential heterogeneities between regions and countries. In addition, since one main focus of our study is to find possible differences in the covariate effects between Asian and African countries, we include a categorical variable that indicates on which continent (Asia, Sub-Saharan Africa) an observation is located. This variable will then be used as an effect modifier to capture potential differences in the determinants between Asian and Sub-Saharan countries through varying coefficients.⁷

Household-specific characteristics. A household's material wealth is positively related with the health status of the children living in this household Fink et al. (2017); Vollmer et al. (2017). To measure the wealth of a household, an asset index is included into the analysis. This index is derived through a principal components analysis based on assets (e.g. television, fridge or motorbike) possessed by the household of observation i and dwelling and sanitation characteristics of the household. For example, Wang (2003) finds that having access to flush toilets is reflected in declining mortality rates.⁸ The effect of the household's wealth enters the hierarchical model in the first level as deviation from the regional specific mean, and it should be an approximation for the household's long term economic wealth. In addition, the regional mean of the asset index is included in the regional specific level of the hierarchical estimation. To capture the effect of the size of the household, the total number of household members is included in the model. The effect is ambiguous as, on the one hand, in a larger household more people potentially can take care of young children, while on the other hand, the conflict about the distribution of the household's resources becomes bigger.

[Please insert Figure 2 about here]

[Please insert Table 2 about here]

⁷Hastie and Tibshirani (1993) first introduced this model where the effect of variable used as effect modifier varies smoothly over the range of the continuous variable.

⁸For a detailed description how the index is derived, see Filmer and Pritchett (2001).

Finally, to account for declining mortality rates in the last two decades (Lozano et al., 2011) and to control for a time trend in under-5 mortality, which potentially declined in a non-linear way, the survey year is included in the estimated models. The right panel of Figure 2 shows the trend in under-5 mortality between 1990 and 2015. Besides a sharp decline in under-5 mortality Figure 2 also reveals considerable level differences in under-5 mortality between Sub-Sahara Africa and Asia, which seem to be stable across time. Besides the afore mentioned individual specific covariates, to account for the clustered data structure, and to estimate the effects of region or country specific covariates we estimate a multilevel model consisting of 3 levels. Table 2 summarized the expected effect of the level specific covariates. At the regional level we include the regional average height for age z-score, and the regional vaccination rate into the analysis. Both variables are unfortunately not directly observed for all sampled children, and therefore including these covariates at the individual level is not possible. We expect the effect of both covariates to be monotonically declining with increasing level of the covariate. As both variables were found to be closely related with under-5 mortality (e.g. Black et al., 2013; Boco, 2010; Charmarbagwala et al., 2004; Wang, 2003). At the country level, the GDP per capita is included into the analysis. A recent study by Baird et al. (2010) found a negative relation between a countries per capita income and under-5 mortality. In particular, a decline of the country’s GDP caused by a recession resulted in an increase in mortality rates. This increase is found to be notably larger for female children (Baird et al., 2010).

2.4. Coding of Covariates

Categorical explanatory variables will be effect coded, since as it is pointed out by Umlauf et al. (2015), the sampling and convergence is improved compared to a dummy variable coding of the explanatory variables. As an example, for the variable *Continent*, which indicates whether a particular observation is located in an Asian or in an Sub-Saharan African country this is done as follows:

$$\text{Continent} = \begin{cases} 1, & \text{if country } i \text{ is located in Sub-Saharan Africa} \\ -1, & \text{if country } i \text{ is located in Asia} \\ 0, & \text{else,} \end{cases}$$

where the corresponding value of the variable takes 1 if country i is located in a Sub-Saharan Africa, and -1 if the country i is located in an Asia, which will be the reference category.

The following covariates enter the model as categorical covariates: whether an older sibling of observation i deceased, the sex of the children, and the place of residence. Additionally, to detect potential differences between Asian and Sub-Saharan African countries a binary indicator is included as the interaction variable in the varying coefficient terms. Descriptive statistics of our sample are presented in Table 3. Around 10% of the sampled children have an older deceased sibling, 70% are born in a household located in living in rural districts, and around half of the sampled children are born as

girls. 35 (73,24%) of the 48 included countries are located in Sub-Saharan Africa.

[Please insert Table 3 about here]

All other variables enter the model as continuous covariates: the deviation from the regional mean of the asset index, the birth order of the child in the household, the years of education of the mother, the numbers of people living in household i , the mother’s age at birth of children i , the mother’s BMI, and to capture a trend over time the year the survey was conducted at the first level. All these variables are found to influence under-5 child mortality by various studies but no study exists that analyzes whether these effects show any non-linearities.⁹

At the second (regional) level, the height for age z-score and the regional immunization rate for specified vaccines are included into the analysis. At the highest level of the hierarchical model (country), we control for the countries average per capita GDP. Unfortunately, information on HIV prevalence and malaria incidence rates is that scarce that we could not include these two covariates into our model. Which is why for countries with high HIV rates and or high malaria incidence under-5 mortality will be underestimated. Table 3 provides some basic descriptive statistics and the level at which the variable enters our baseline model.

3. Methodology

3.1. Discrete survival model

Discrete survival models can be specified as regression models with binary response, for instance as logit or probit model. However, to make an estimation feasible, the data needs to be augmented, such that the number of entries of observation i are equal to the number of intervals t observation i is exposed to the risk of dying. For a more elaborate discussion of the data augmentation consult the methodology manual of **BayesX**.

DHS provides only monthly information on the age of the child at the date of the interview. Hence, information on the duration is only measured on a discrete time scale¹⁰. Duration time T can be considered as discrete with $T \in \{1, 2, \dots, q; q = 10\}$. Besides the duration time T , the vector of covariates x_i^* is observed, then $x_{it}^* = (x_{i1}, x_{i2}, \dots, x_{it})$ denotes the information at each interval up to interval t . Following Adebayo and Fahrmeir (2005), the discrete hazard function can be written as follows:

$$\lambda(t; x_{it}^*) = P(T = t | T \geq t, x_{it}^*), \quad t = 1, \dots, 10, \quad (1)$$

⁹Also birth spacing is found to influence the health outcome of children Mosley and Chen (1984); Rutstein (2005). However, this information is not available for first born children, and omitting this group would systematically bias our results. Hence, the final model is re-estimated including information on children i ’s preceding birth interval.

¹⁰Additionally, due to restricted memory, and to make an estimation tractable in **BayesX**, we further divide the duration time scale into semiannual intervals, each of length of six months, $[0, 1, 2, 3, 4, 5], [6, 7, 8, 9, 10, 11], \dots, [54, 55, 56, 57, 58, 59]$.

where $\lambda(t; x_{it}^*)$ is the discrete hazard function rewritten as the conditional probability of death in interval t , given child i survived until interval t .

The event indicator $dead_{it}$, with $dead_{it} = 0$ if child i is alive in interval t , and $dead_{it} = 1$ if child i died in interval t . In the former case, t_i corresponds to the current semiannual age interval of child i at the interview, and in the latter case t_i corresponds to the observed age interval the child died. Thus the binary event indicator $dead_{it}$ of the survival status at interval t_i for child i is defined as follows:

$$dead_{it} = \begin{cases} 1, & \text{if } t = t_i, \text{ and } dead_i = 1 \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Thus, the underlying survival process of child i can be interpreted as series of dichotomous choices at each interval t . Assuming a probit link for the discrete-time survival model, the response function is given by $\pi_{it} = \phi(\eta_{it})$, where $\phi(\eta_{it})$ is the cumulative distribution function of the standard normal distribution and the link function is given by $\phi^{-1}(\pi_{it}) = \eta_{it}$ (Fahrmeir et al., 2013, Chapter 4). The linear predictor for under-5 mortality is given by η_{it} , and, as in Adebayo and Fahrmeir (2005), the partially linear predictor can be written as follows:

$$\eta_i^{lin} = \gamma_0(t) + \gamma_1 x_{it1} + \dots + \gamma_q x_{itq} + \epsilon_{it}. \quad (3)$$

Using only the partially linear predictor η_{it}^{lin} leads to the restriction that the relationships of continuous covariates have to be linear, which is not justified for all covariates. Most studies that have analyzed determinants of child health in general and under-5 mortality in particular have followed such an approach. In view of this pitfall, in our study, the linear predictor η_{it}^{lin} is substituted by the more general predictor η_{it} , in which the continuous covariates are modeled through a potentially non-linear function f . This yields to the following additive model:

$$\eta_{it} = \mathbf{x}'_{it} \boldsymbol{\gamma} + f_0(t) + f_1(x_{it1}) + \dots + f_p(x_{itp}) + \epsilon_{it}, \quad (4)$$

where $\mathbf{x}'_{it} \boldsymbol{\gamma}$ is the set of covariates that enter the model linearly (the effect coded categorical covariates), $f_0(t)$ captures the baseline effect, and $f(x_{it1})$ to $f(x_{itp})$ are continuous covariates that enters the model with an assumed non-linear relationship (Lang et al., 2014).

3.2. Modelling the non-linear function f

The potential non-linear terms will be modeled through a non-linear function f that can be estimated using P-splines (Fahrmeir et al., 2013, Chapter 7). In this setting, the penalized splines will be approximated by using an adequate linear combination of basis

functions that take the form:

$$f(x) = \sum_{j=1}^d \gamma_j B_j(x) \quad (5)$$

The basic idea of P-splines is to divide the range of the data into a relatively large number of intervals. The boundaries between two intervals are called knots, and we use ten knots as the default choice for the number of knots. At the same time a penalty term is included to prevent overfitting. In a frequentist setup this model could be estimated considering standard maximum likelihood procedure. In a Bayesian setup the model will be estimated using a simulation procedure through MCMC.¹¹

3.3. Capturing differences between Asian and Sub-Saharan African countries with varying coefficients

For determining potential differences between Asian and Sub-Saharan African countries a varying coefficient term for the geographic location is included in the estimation. Including varying coefficients was first proposed by Hastie and Tibshirani (1993) and applications of varying coefficients can be found in Lang and Sunder (2003) or Lang et al. (2014), where in the second paper Indian data is used to estimate determinants of children’s malnutrition with respect to potential gender differences. In this context, the effect of a covariate, e.g. the BMI of the mother can vary with respect to binary indicator for Asian respectively Sub-Saharan African countries. This will be modeled as follows,

$$\eta_{it}^{add} = \dots + f_{12}(\text{BMI mother}_{it}) + f_{\text{BMI mother}|\text{Continent}}(\text{BMI mother}_{it})(\text{Continent}) + \dots + \mathbf{x}'_{it}\boldsymbol{\gamma} + \epsilon_{it}.$$

where the categorical covariate *Continent* is the interaction term and the continuous covariate *BMI* the effect modifier. The varying coefficients term now allows the effect of of the region (Asia versus Sub-Sahara Africa) to vary smoothly across the range of the mother’s BMI. Including this varying coefficient into the estimation should detect potential differences in the effect of a specific covariate for Asian and African countries.

3.4. Multilevel STAR framework

To account for the hierarchical structure of the data, and to allow for heterogeneity between countries, a multilevel STAR model is used. The 824,017 observed children are nested into 358 regions, which are again nested into 48 countries in Asia and Sub-Saharan Africa, this is nesting of the observations is also called clustering (Jain et al., 1999; Korenromp et al., 2004). Following Lang et al. (2014), who developed a multilevel

¹¹For a more thorough description of the estimation see for instance (Fahrmeir et al., 2013, Chapter 7 and Chapter 8) or Harttgen et al. (2015).

approach that allows to account for heterogeneity at the regional and country level, as well as to include non-linear covariate effects. The multilevel model with 3 levels and which allows to account for the afore mentioned structure is written as follows:

$$\begin{aligned}
\text{Level - 1 : } \eta_{it} &= \eta_{it}^{\text{Individual}} + f_{18}(\text{Region}_{it}) + \epsilon_{it} \\
\text{Level - 2 : } f_{18}(\text{Region}) &= f_{18,1}(\text{Stunting regional mean}_{\text{Region}}) + f_{18,2}(\text{Stunting regional mean}_{\text{Region}})\text{Continent} + f_{18,3}(\text{Regional vaccination rate}_{\text{Region}}) + f_{18,4}(\text{Regional vaccination rate}_{\text{Region}})\text{Continent} + f_{18,5}(\text{Country}_{\text{Region}}) + f_{18,6}(\text{Region}) + \epsilon_{18} \\
\text{Level - 3 : } f_{18,5}(\text{Country}) &= f_{18,5,1}(GDP_{\text{Country}}) + f_{18,5,2}(GDP_{\text{Country}})\text{Continent} + f_{18,5,3}(\text{Country}) + \gamma_{18,5,1}\text{Continent}_{\text{Country}} + \epsilon_{20,5}
\end{aligned} \tag{6}$$

This model accounts for the hierarchical structure of the data, which allows to model the effects of regional, respectively country specific covariates on under-5 mortality. To determine systematic differences between Asian and Sub-Saharan African countries a varying coefficient is included in the STAR model. In Equation 6, $\eta_{it}^{\text{Individual}}$ subsumes the included categorial covariates, and the continuous covariates including the corresponding varying coefficient, see Table 2 for the included covariates and the expected effects.

Accounting for heterogeneity Up to this point the baseline model captures cluster specific heterogeneity only by including i.i.d. Gaussian random effects of the region ($f_{18,6}(\text{Region}) \sim \mathcal{N}(0; \tilde{\sigma}_{18,6}^2)$), and the country ($f_{18,5,3}(\text{Country}) \sim \mathcal{N}(0; \tilde{\sigma}_{18,5,3}^2)$). To capture potential heterogeneity of the individual specific non-linear effects f_q , and the individual specific linear effects γ_q , for each individual specific covariate a country specific random effect is included.

Models that include this type of interaction for an unordered grouping factor, such as countries, are also called models with random slopes (Belitz et al., 2015). This can be of particular use, when accounting for country or region specific heterogeneity. The basic idea is then to add random slopes to the individual specific covariates of Equation 6. For example, for the effect of the BMI of the mother, the model is rewritten as follows:

$$\begin{aligned}
\eta_{it} = \dots + f_{12}(\text{BMI mother}_{it}) + \tilde{f}_{12}(\text{Country})(\text{BMI}_{it}) + f_{13}(\text{BMI mother}_{it})(\text{Continent}) + \dots + \epsilon_{it}.
\end{aligned} \tag{7}$$

Where $\tilde{f}_{12}(\text{Country}) \sim \mathcal{N}(0; \tilde{\sigma}_{11}^2)$ is i.i.d. normal distributed with inverse Gamma distributed variance $\tilde{\sigma}_{12}^2$. A more recent, alternative method to model the cluster specific variation of the non-linear terms is to use multiplicative country specific random effects, as they were applied in recent papers, see for instance Wechselberger et al. (2008); Brunauer et al. (2013); Lang et al. (2015). In this variant the country specific random

effect is written as follows,

$$\eta_{it} = \dots + (1 + \tilde{f}_{12}(\text{country}_{it})) \cdot f_{12}(\text{BMI mother}_{it}) + f_{13}(\text{BMI mother}_{it})(\text{Continent}) + \dots + \epsilon_{it}. \quad (8)$$

A random effect smaller than zero ($\tilde{f}_q(\text{country}_i) < 0$) yields to a scaling down of the main effect, whereas a random effect greater than zero ($\tilde{f}_q(\text{country}_i) > 0$) yields the main effect to be scaled up. For another application of multiplicative country specific random effects see Harttgen et al. (2016).

3.5. Model selection

Simultaneous Bayesian credible bands Krivobokova et al. (2010) together with the deviance information criterion (DIC) Spiegelhalter et al. (2002) are used for model selection and analyzing significance of the covariates. The DIC is considered to be the generalization of the AIC and the BIC for hierarchical models. Moreover, it is pointed out by Spiegelhalter et al. (2014) the AIC can not be calculated in hierarchical models, whereas it is feasible to calculate the DIC. As Bayesian models became feasible in the late 1990s a model comparison criterion was still missing, and the DIC was proposed by Spiegelhalter et al. (2002) in 2002. A more detailed description of the selection process can be found on Appendix B.

4. Results

4.1. Linear effects

Table 4 summarizes the results of the linear covariates. The table shows the posterior mean together with the 95% credible interval. The only effect found to be insignificant, was whether the household is lead by a female or male in the model building phase and was, therefore, not included into the final analysis. All other linear covariates are found to be significant¹². Beginning with the covariate showing the largest effect, the indicator of whether the child has a older sibling who have died or not. Growing up in a household where an older sibling died is associated with a higher mortality risk in the first 60 months, compared to a household where no such fatality occurred. The estimated posterior mean is 0.337 which is in line with our expectations, as well as the literature, see for instance Hobcraft et al. (1985); Harttgen et al. (2016).

[Please insert Table 4 about here]

The estimated posterior mean of 0.028 of the gender of the child is also significant and in line with our expectations. Boys tend to have a slightly higher risk of dying in the first five years of living, compared to girls. This can also be illustrated by plotting the

¹²A linear effect is assumed to be significant if zero is not included in the 95% credible interval.

Nelson-Aalen cumulative hazard for the individual waves. In the left hand panel of Figure 2 the cumulative hazard rate is plotted separately for male and female observations from Asia, and for male and female observations Sub-Saharan African countries. It can be seen that the hazard rates for girls is always slightly below the hazard rate for boys.

Children living in urban areas, have higher survival probabilities compared to their counterparts in rural areas. The estimated posterior mean of -0.018 is significant and negative, however of low magnitude. This estimates reveal that living in rural areas reflects that children below five years are confronted with a higher mortality risk. This gap between urban and rural areas might be best explained through generally poorer health status in rural areas and a poorer basic health infrastructure in rural areas (Harttgen and Günther, 2011). However, the estimated results for the gender of the the child, and the area of living should be interpreted with great care. Especially when large sample sizes are analyzed already relatively small effects tend to become significant. Both effects are significant but scarcely important, considering the usual predictor range of probit models between -2 and 2 .

It can be assumed that the effect size of the categorical covariates is not homogeneous across countries, and we included country random slopes for the categorical covariates. Figure 3 and Figure 4 shows the results of the random slopes of the categorical covariates. Especially, for the variables whether the child had a sibling who died, the place of residence, and the children's gender considerably differences across countries exist.

[Please insert Figure 3 about here]

[Please insert Figure 4 about here]

4.2. Non-linear effects

The estimated STAR model simultaneously takes into account possible non-linearities, and the hierarchical structure of the data. This is especially useful, since some covariates can not be observed on an individual basis and only regional or country aggregates can be included. Beginning with the effects at the individual level, the results are discussed in the next section.

Individual level Figure 5 to the top panel of Figure 12 show the results of the non-linear effects at level 1 of the individual specific continuous covariates. The figures show from left to right, the mean effect together with the 95% and 80% credible intervals, the continent specific effect for Asia and Sub-Saharan Africa, and the continent specific effects scaled by the country random effect for the continuous covariates. In more detail for the baseline hazard: The top-left panel in Figure 5 depicts the posterior mean of the mean effect including credible intervals. The mortality risk is found to be highest in the first month, and than steadily declining. Two things become obvious in this Figure: First, the mortality risk is highest in the first months, and the baseline hazard having a concave shape. Second, the sharp decline around the 50th month seems to be an artifact, a similar pattern was already pointed out by Adebayo and Fahrmeir (2005).

Most likely, for older children the age at death is only reported in full years, instead of a more accurate monthly reporting. Besides the baseline effect, the varying effect (see left panel Figure C.1) is found to be significant. These two effects can be combined to the top right panel of Figure 5, which illustrates mayor differences between Asian and Sub-Saharan African countries. In particular, a faster decrease of the probability of dying for observations located in Asia. Relatively to the effect of Asian countries, for Sub-Saharan African countries a flatter baseline hazard is observed, indicating a mortality risk that is less depending on the age of the child. Modelling country specific heterogeneities with multiplicative random effects yields substantial differences between countries, and to a scaling down of the baseline hazard when the random effect is < 0 and to a scaling up of the baseline effect if the random effect is > 0 . Combining the country specific random effect (see right panel Figure C.1) and the continent specific effect (bottom left panel Figure 5) gives the county specific effects illustrated in the two bottom panels of Figure 5. Interpreting the results of the other covariate effects is analogous.

[Please insert Figure 5 about here]

The effect of the age of the mother at birth has its presumptive inverse J-/U-shaped form. In particular, children of young mothers are faced with a higher mortality in the first five years, decreasing relatively fast and flattening out for mothers older than 20 years (Figure 10). This result is consistent with findings of Harttgen et al. (2015), who analyzed determinants of child mortality for a Sub-Saharan African sample, and with a way earlier work by Hobcraft et al. (1985). The mortality risk declines sharply up to an age of the mother at birth of around 20 years. The effect then flattens out, remains more or less constant for women up to an age around 30 years, before it slightly increases for woman older 40 years. Even though the with the continent varying effect is significant, it is hardly relevant due its magnitude in comparison to the magnitude of the main effect.

[Please insert Figure 10 about here]

Figure 7 depicts the effect of the households birth order and reveals an interesting finding. The effect is monotonically increasing over the complete range, the presumptive U-shaped effect which was already found by Hobcraft et al. (1985) was not found. Being higher in the birth order, reflects a rising mortality risk. Also the varying effect is significant, resulting in a effect for Sub-Saharan African countries that is less pronounced as for Asian countries, and the birth order has a stronger effect on under-5 mortality in Asian countries. Across countries to some extent differences exist, however not of the same magnitude found, for instance, for the household's wealth or the size of the household, and differences are only modest.

[Please insert Figure 7 about here]

The size of the household is also found to influence under-5 mortality significantly with children's living in small households ($n \leq 3$) are found to be faced with the highest mortality risk (Figure 9). Most likely this corresponds with the effect of growing up in a single parent household, and only one parent is able to provide income and food. The effect decreases monotonically until a household size of approximately 7 or 8 people. For larger households the effect remains more or less constant. The varying effect of the household size is found to be significant, however irrelevant due to the magnitude, and differences between households located in Asian countries or Sub-Sahara African countries can be ignored.

[Please insert Figure 9 about here]

Contrary to Harttgen et al. (2015), their finding of a inverted U-shaped relationship between the mother's BMI and child mortality is reversed and the effect is U-shaped (Figure 11). It decreases up to a BMI of around 20, and then increases again. This is particularly interesting since malnutrition of the mother and also obesity of the mother increase child mortality. However the effect is found to be rather small in magnitude. Not surprising the effect is rather similar for Asian countries and Sub-Saharan countries, a little less dominant in Sub-Saharan African countries. Additionally the effect is rather homogeneous across countries, showing only little heterogeneity.

[Please insert Figure 11 about here]

The effect of the household's wealth is decreasing more or less monotonically (Figure 6). This finding is of particular interest since reducing poverty rates in developing countries would also imply that child mortality will decline in these countries. The interaction with geographic indicator (continent) is literally zero and reveals that no differences exist between households located in Asian countries and households located in Sub-Saharan African countries. However a large dispersion exists between countries and the effect shows in particular for Asian countries a large variation.

[Please insert Figure 6 about here]

Another rather strong effect which is noteworthy is the effect of the mother's years of education (Figure 8). The effect is found to be more or less decreasing linearly, and corresponds to the findings of Cleland and van Ginneken (1988). Investing in the education for expecting mothers, and women in general, seems to lower child mortality. This can reflect 2 major thoughts. First, women with higher education can better assess relevant information which, in the end, will benefit the children's health and overall status. Second, higher education has a monetary payoff, since the chance to find better jobs increases. However one should also keep in mind the explanation of Desai and Alva (1998), who argue that the educational level of the mother can be seen as an approximation of the family's socioeconomic status.

[Please insert Figure 8 about here]

Declining under-5 mortality rates are observed since the late 1990s and can be illustrated by plotting under-5 mortality rates by year (see right hand panel of Figure 2), and are found by various authors (see for instance Lozano et al., 2011; Jamison et al., 2016). We find that under-5 mortality in Asian countries was decreasing constantly over time (Figure 12). Whereas, in comparison to Asian countries for Sub-Saharan African countries a significant difference could be found. This effect of Sub-Saharan African countries is almost constant for the time before 2000 and only after 2000 mortality rates started to decline. This is consistent with a study by Black et al. (2003), where Sub-Saharan African countries experienced a drop in under-5 mortality too, however the decline was of much smaller magnitude in comparison to Asian countries.

[Please insert Figure 12 about here]

Regional and country level The estimated model includes the regional mean of the height for age z-score, the composite regional immunization rate for selected vaccines, and at the country level the country mean of the GDP. Only the main effects are displayed since no varying effect was found to be significant. The effects are depicted in the bottom panel of Figure 12. A better nutritional status of a particular region reduces the individual mortality risk for observations in this region. The effect is more or less constant for values below -2 (the threshold of the WHO for moderate stunting) and then monotonically decreasing. Differences between Asian and Sub-Saharan countries are found not to be significant, and therefore omitted in the final model. This effect is consistent with the general findings of Pelletier and Frongillo (2003); Black et al. (2003, 2013), however puts an emphasis on reducing malnutrition (measured as the height for age z-score) below the threshold of -2 of the WHO for moderate stunting.

Considering the regional immunization rate, the effect is more or less monotonically decreasing. Showing that in regions with lower immunization rates children are exposed to a higher risk of dying. These findings can be directly transferred to the advice for policy makers and institutions that reducing the number of severely stunted children, and increasing the share of children fully vaccinated against BCG, measles, DPT, and polio would also be reflected in decreasing mortality rates. This finding emphasizes the need for better vaccination campaigns to increase immunization coverage as already recommended lately by Boco (2010); Wang (2003). In the case of vaccination coverage, institutions should aim for a higher vaccination rate across countries, that follows closely the guidelines established by the WHO.

At the country level, the effect of the countries per capita GDP is pointing in the right direction, however is found to be insignificant. The effect is more or less monotonically decreasing showing no differences between Asian and Sub-Saharan African countries. This contradicts at least partially the findings of Baird et al. (2010) that solely increasing GDP per capita does not effectively reduce under-5 mortality, and as mentioned by Wang (2003) an emphasis should be put on improving a country's health system. However, more research is necessary to establish a better understanding how an increase in the per capita GDP can be transferred to reduced under-5 mortality rates.

4.3. Random effects

To account for heterogeneities across countries, we include country specific random slopes for the categorial covariates, and rely on a multiplicative random effects for the non-linear terms. Figure 3 and Figure 4 show the results of the country specific effect of the linear terms. These Figures depict the posterior means together with the 80% and 95% simultaneous credible intervals, and reflect the existing huge dispersion between countries with respect to the categorial covariates.

Country specific random effects for the level-1 non-linear covariates are shown in the bottom-right panels of Figure C.1 to Figure C.7 for each covariate separately. As mentioned, among others, by Harttgen et al. (2016) multiplicative random effects cause a scaling of the estimated main effect, where a positive random effect yields to an up scaling of the main effect, and a negative random effect yields to a down scaling of the main effect. Differences between countries exist especially for the effect of the asset index, the household size, and the age of the mother at birth. Heterogeneity across countries for the other effects is less pronounced, however still existing for the BMI of the mother, the birth order and the level of education of the mother.

4.4. Interaction effects

To control, whether systematic differences between the two analyzed regions, with respect to the children's gender and the children's place of living exist, we re-estimate the main model including interaction terms for the two mentioned categorial covariates with the geopolitical location. This will identify, if potential differences exist, and when differences exist it can help to formulate policies that aim to reduce child mortality.

Controlling for a possible interaction between the child's gender and the geographic location of the household confirms the pattern already illustrated in the left hand panel of Figure 2. Being born as boy within Sub-Saharan Africa is associated with the highest mortality risk (posterior mean: 0.149; 80% credible interval: [0.059; 0.200]), followed by being born as girl within Sub-Saharan Africa (0.092; [0.003; 0.110]). Since categorial covariates are effect coded, it is possible to derive an estimate for the reference category (being born as girl within Asia), the posterior mean is -0.130 and almost identical to the estimate of begin born as boy within Asia (-0.111; [-0.161; -0.019]).

Allowing for a similar interaction, however this time for the household's place of living and the geographic location, reveals similar results as described in the 'baseline' model. Children born in urban areas, possibly due to better access to health care, sanitation, and drinking water, are exposed to a lower mortality risk compared to children growing up in rural areas. Table 5 shows these results in more detail.

[Please insert Table 5 about here]

Determining whether differences between Central Asia, South Asia and Sub-Saharan Africa exist, reveal a few interesting patterns. Additionally, distinguishing Asian countries between Central Asian countries and South Asian countries improves the fit of the

model by roughly 350 points. Removing insignificant varying coefficient terms, and only including those which were found to be significant does not change the DIC much.

Accounting for differences between Central Asia and South Asia was found to be necessary for the baseline hazard, the birth order, the household size, and the age of the mother at birth. Results are shown in the Appendix in Figure D.1 to Figure D.8. The shape of the effects are rather similar to the effects of the baseline model, however the magnitude of the effects differs between the geopolitical regions.

5. Conclusion and Discussion

On a global scale, since 1990 under-5 mortality fell approximately by one half, moreover a decline both in relative terms (per 1,000 live-birth), as well as in absolute numbers is observed (You, Hug, Ejdemyr, Idele, Hogan, Mathers, Gerland, New and Alkema, 2015). Despite this positive development, two regions can be identified where the reduction was lacking behind the projections needed to achieve the MDG 4 of reducing under-5 mortality by two thirds between 1990 and 2015, these regions are: Asia (in particular Caucasus and Central Asia, and Southern Asia) and Sub-Saharan Africa (You, New and Wardlaw, 2015). These regions showed insufficient progress to achieve the target of the MDG. Analyzing jointly these regions¹³ will give new insights in the determinants of under-5 mortality. This is thought to offer a starting-point to improve policy towards achieving the target of SDG 3: reducing under-5 mortality to at least as low as 25 per 1,000 live births, by 2030 (Liu et al., 2016).

For our analysis, we take advantage of the structure, standardization, and comparability of the DHS data sets, which allows a comparison across countries and over time. We jointly analyze under-5 mortality for two geographic regions: Sub-Sahara Africa and Asia, considering in total 48 countries in 131 individual surveys over the period from 1992 to 2015. In total, we analyze 824,017 children, using a new multilevel discrete time survival model, that takes censoring into account, and observations are clustered into 358 distinct regions, and 48 countries. Applying the hierarchical STAR framework developed by Lang et al. (2014), allows us to take non-linear covariate effects into account, estimate differences between Asian and Sub-Saharan African countries through varying coefficients (Hastie and Tibshirani, 1993), and allows us to account for cluster specific heterogeneity. Estimation of this type of hierarchical models with structured additive predictor is fully Bayesian, and non-linear covariate effects are estimated through Bayesian P-splines.

Following the conceptual framework to study child health in low and lower-middle income countries proposed by Mosley and Chen (1984), we analyze determinants of under-5 mortality in four dimension of factors associated with child wellbeing: child-specific factors, maternal factors, environmental and socio-economic factors, and household specific factors.

However, also this study has some limitations which should be recognized. First, DHS

¹³The DIC do not provide sufficient data for countries located in Oceania, therefore this region was no part of the analysis.

do not observe information of important determinants for deceased children, implying that this variables could not be included into our estimation. For instance, Rutstein (2000) found that there is a correlation between breastfeeding and mortality. However this relationship, due to missing information for deceased children could not be observed. Additionally, questionnaires are prone to problems in the data collecting process. Missing values can be introduced in the data, as the respondent fails to answer particular question. For instance the birth weight, which shows a large number of missing values could not be remembered by the respondent, or fatal incidences are under reported within sampled household. Second, one of the main underlying causes of under-5 mortality found by several authors (e.g. Pelletier and Frongillo, 2003; Black et al., 2003, 2013). Unfortunately, DHS provides information about the nutritional status only for those children who where still living at the point in time the interview was conducted. This allows us to only include the regional mean of the normalized anthropometric indicator into the level-2 equation of the hierarchical model. By doing so, this gives rise to a potential sampling error, namely so called selective mortality Alderman et al. (2011). Only including the regional mean of the anthropometric indicator into the model gives rise to a potential bias. In the work by Alderman et al. (2011) this issue was addressed for three waves of Indian national health surveys, and no significant change was found. Analyzing this aspect in more detail on a global scale, using the available DHS data will offer better insights, whether this caveat is of substantial nature, and whether the relationship between malnutrition, and under-5 mortality is underestimated at a significant level.

The major strength of the analysis is that, for the first time, determinants of under-5 mortality in the geographic regions that where found to fall short in achieving the MDG 4 are analyzed jointly, using a multilevel discrete time survival model. Potential differences between these two regions are identified, and taking advantage of a large sample with 824,017 observations, we account for non-linear covariate effects on under-5 mortality through Bayesian P-splines. Additionally, we use a recently developed multilevel framework in a Bayesian setting that allows us to account for country and region specific heterogeneity.

Assuming only linear relationships between under-5 mortality and its determinants is misleading in the sense that the effect is independent of the determinants level. We find strong non-linear effects for this relationships, that can be summarized as follows. Differences between Asia and Sub-Saharan Africa exist with respect to the baseline hazard, the household's wealth, the household size, the surveyyear, and of minor extent for the birth order. We find relatively strong non-linear effects for the baseline hazard, the birth order, the household size, and the age of the mother at birth. Additionally, the effect of the mother's nutritional situation is found to have a U-shaped form, however the effect is less dominant. Minor non-linearities are found for the household's wealth, and the years of education of the mother. Considering the regional level, both covariates are found to be significant. Improving the nutritional situation, as well as increasing immunization coverage will effectively reduce mortality. No significant impact was found at the country level for the per capita GDP, however the effect is pointing in the right direction.

No single factor could be isolated on which policy should focus on, instead our analysis shows that a wide range of influential determinants on under-5 mortality exists. However, improving maternal education, immunization rates, and lowering the number of malnourished children can be seen as starting points to which policy should be directed. Further research needs to control for HIV and malaria prevalence, however this requires a significant improvement of data availability. Additionally, estimating spatial effects and identifying high risk areas could improve policy aiming to reduce under-5 mortality.

References

- Adebayo, S. B. and Fahrmeir, L. (2005). Analysing child mortality in nigeria with geoaddivitive discrete-time survival models, *Statistics in Medicine* **24**(5): 709–728.
URL: <http://dx.doi.org/10.1002/sim.1842>
- Alderman, H., Lokshin, M. and Radyakin, S. (2011). Tall claims: Mortality selection and the height of children in India, *Economics & Human Biology* **9**(4): 393 – 406.
URL: <http://www.sciencedirect.com/science/article/pii/S1570677X11000487>
- Ayele, D. G. and Zewotir, T. T. (2016). Childhood mortality spatial distribution in ethiopia, *Journal of Applied Statistics* **43**(15): 2813–2828.
URL: <http://dx.doi.org/10.1080/02664763.2016.1144727>
- Ayele, D. G., Zewotir, T. T. and Mwambi, H. G. (2015). Structured additive regression models with spatial correlation to estimate under-five mortality risk factors in Ethiopia, *BMC Public Health* **15**(1): 268.
URL: <http://dx.doi.org/10.1186/s12889-015-1602-z>
- Baird, S., Friedman, J. and Schady, N. (2010). Aggregate Income Shocks and Infant Mortality in the Developing World, *Review of Economics and Statistics* **93**(3): 847–856.
URL: http://www.mitpressjournals.org/doi/abs/10.1162/REST_a.00084
- Belitz, C., Brezger, A., Klein, N., Kneib, T., Lang, S. and Umlauf, N. (2015). *BayesX: Software for Bayesian Inference in Structured Additive Regression Models. Version 3.0.2.*
URL: <http://www.BayesX.org/>
- Bhalotra, S. (2006). Childhood Mortality and Economic Growth, *Working Paper Series RP2006/79*, World Institute for Development Economic Research (UNU-WIDER).
URL: <https://ideas.repec.org/p/unu/wpaper/rp2006-79.html>
- Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., Mathers, C. and Rivera, J. (2008). Maternal and child undernutrition: global and regional exposures and health consequences, *The Lancet* **371**(9608): 243 – 260.
URL: <http://www.sciencedirect.com/science/article/pii/S0140673607616900>
- Black, R. E., Morris, S. S. and Bryce, J. (2003). Where and why are 10 million children dying every year? , *The Lancet* **361**(9376): 2226 – 2234.
URL: <http://www.sciencedirect.com/science/article/pii/S0140673603137798>
- Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., Martorell, R. and Uauy, R. (2013). Maternal and child under-nutrition and overweight in low-income and middle-income countries, *The Lancet* **382**(9890): 427 – 451.
URL: <http://www.sciencedirect.com/science/article/pii/S014067361360937X>
- Boco, A. G. (2010). Individual and community level effects on child mortality: An analysis of 28 Demographic and Health Surveys in Sub-Saharan Africa, *Technical report*, Macro International, Calverton, Maryland, USA.
URL: <http://dhsprogram.com/pubs/pdf/WP73/WP73.pdf>
- Brezger, A., Kneib, T. and Lang, S. (2005). BayesX: Analyzing Bayesian Structural Additive Regression

- Models, *Journal of Statistical Software* **14**(1): 1–22.
URL: <https://www.jstatsoft.org/index.php/jss/article/view/v014i11>
- Brunauer, W., Lang, S. and Umlauf, N. (2013). Modelling house prices using multilevel structured additive regression, *Statistical Modelling* **13**(2): 95–123.
URL: <http://smj.sagepub.com/content/13/2/95.abstract>
- Chamarbagwala, R. (2010). Sibling composition and selective gender-based survival bias, *Journal of Population Economics* **24**(3): 935–955.
URL: <http://dx.doi.org/10.1007/s00148-010-0314-z>
- Charmarbagwala, R., Ranger, M., Waddington, H. and White, H. (2004). The determinants of child health and nutrition: a meta-analysis, *Technical report*, Washington, DC : World Bank Group.
URL: <http://documents.worldbank.org/curated/en/505081468327413982/The-determinants-of-child-health-and-nutrition-a-meta-analysis>
- Cleland, J. G. and van Ginneken, J. K. (1988). Maternal education and child survival in developing countries: The search for pathways of influence, *Social Science & Medicine* **27**(12): 1357 – 1368.
URL: <http://www.sciencedirect.com/science/article/pii/0277953688902018>
- de Onis, M., Blössner, M. and Borghi, E. (2012). Prevalence and trends of stunting among pre-school children, 1990–2020, *Public Health Nutrition* **15**: 142–148.
URL: http://journals.cambridge.org/article_S1368980011001315
- Deaton, A. (2007). Height, Health, and Development, *Proceedings of the National Academy of Sciences of the United States of America* **104**(33): 13232–13237.
URL: <http://www.jstor.org/stable/25436467>
- Desai, S. and Alva, S. (1998). Maternal education and child health: Is there a strong causal relationship?, *Demography* **35**(1): 71–81.
URL: <https://doi.org/10.2307/3004028>
- Fahrmeir, L., Kneib, T., Lang, S. and Marx, B. (2013). *Regression: Models, Methods and Applications*, 1 edn, Springer-Verlag Berlin Heidelberg.
- Feenstra, R. C., Inklaar, R. and Timmer, M. P. (2015). The Next Generation of the Penn World Table, *American Economic Review* **105**(10): 3150–82.
URL: <http://www.aeaweb.org/articles?id=10.1257/aer.20130954>
- Filmer, D. and Pritchett, L. H. (2001). Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application To Educational Enrollments In States Of India, *Demography* **38**(1): 115–132.
URL: <http://dx.doi.org/10.1353/dem.2001.0003>
- Fink, G., Victora, C. G., Harttgen, K., Vollmer, S., VIDALETTI, L. P. and Barros, A. J. (2017). Measuring socioeconomic inequalities with predicted absolute incomes rather than wealth quintiles: A comparative assessment using child stunting data from national surveys, *American Journal of Public Health* **107**(4): 550–555. PMID: 28207339.
URL: <http://dx.doi.org/10.2105/AJPH.2017.303657>
- Harttgen, K. and Günther, I. (2011). Deadly Cities? A Note on Spatial Inequalities in Mortality in Sub-Saharan Africa, *Technical Report 39*, ETH Zurich, Kiel und Hamburg.
URL: <http://hdl.handle.net/10419/48302>
- Harttgen, K., Lang, S. and Santer, J. (2015). Multilevel modelling of child mortality in Africa, *Technical report*, Working Papers in Economics and Statistics.
URL: <ftp://ftp.repec.org/opt/ReDIF/RePEc/inn/wpaper/2015-03.pdf>
- Harttgen, K., Lang, S. and Santer, J. (2016). A multilevel approach to estimate non-linear effects in determinants of under-5 mortality in Sub-Saharan Africa, *Technical report*, Working Papers in Economics and Statistics University of Innsbruck.
- Hastie, T. and Tibshirani, R. (1993). Varying-Coefficient Models, *Journal of the Royal Statistical Society. Series B (Methodological)* **55**(4): 757–796.
URL: <http://www.jstor.org/stable/2345993>
- Hobcraft, J., McDonald, J. and Rutstein, S. (1985). Demographic Determinants of Infant and Early Child Mortality: A Comparative Analysis, *Population Studies* **39**(3): 363–385.
URL: <http://dx.doi.org/10.1080/0032472031000141576>
- Houweling, T. A., Kunst, A. E., Looman, C. W. and Mackenbach, J. P. (2005). Determinants of

- under-5 mortality among the poor and the rich: a cross-national analysis of 43 developing countries, *International Journal of Epidemiology* **34**(6): 1257–1265.
URL: <http://ije.oxfordjournals.org/content/34/6/1257.abstract>
- Jain, A. K., Murty, M. N. and Flynn, P. J. (1999). Data clustering: A review, *ACM Comput. Surv.* **31**(3): 264–323.
URL: <http://doi.acm.org/10.1145/331499.331504>
- Jamison, D. T., Murphy, S. M. and Sandbu, M. E. (2016). Why has under-5 mortality decreased at such different rates in different countries?, *Journal of Health Economics* **48**: 16 – 25.
URL: <http://www.sciencedirect.com/science/article/pii/S0167629616300029>
- Kandala, N.-B. and Ghilagaber, G. (2006). A Geo-Additive Bayesian Discrete-Time Survival Model and its Application to Spatial Analysis of Childhood Mortality in Malawi, *Quality and Quantity* **40**(6): 935–957.
URL: <http://dx.doi.org/10.1007/s11135-005-3268-6>
- Kandala, N.-B., Mandungu, T. P., Mbela, K., Nzita, K. P., Kalambayi, B. B., Kayembe, K. P. and Emina, J. B. O. (2014). Child mortality in the democratic republic of congo: cross-sectional evidence of the effect of geographic location and prolonged conflict from a national household survey, *BMC Public Health* **14**(1): 266.
URL: <http://dx.doi.org/10.1186/1471-2458-14-266>
- Klasen, S. (2000). Malnourished and surviving in South Asia, better nourished and dying young in Africa: what can explain this puzzle?, *Technical report*, University of Munich.
- Korenromp, E., Arnold, F., Williams, B., Nahlen, B. and Snow, R. (2004). Monitoring trends in under-5 mortality rates through national birth history surveys, *International Journal of Epidemiology* **33**(6): 1293–1301.
URL: <http://ije.oxfordjournals.org/content/33/6/1293.abstract>
- Krivobokova, T., Kneib, T. and Claeskens, G. (2010). Simultaneous Confidence Bands for Penalized Spline Estimators, *Journal of the American Statistical Association* **105**(490): 852–863.
URL: <http://dx.doi.org/10.1198/jasa.2010.tm09165>
- Kuate-Defo, B. and Diallo, K. (2002). Geography of child mortality clustering within african families, *Health & Place* **8**(2): 93 – 117.
URL: <http://www.sciencedirect.com/science/article/pii/S1353829201000387>
- Lang, S., Steiner, W. J., Weber, A. and Wechselberger, P. (2015). Accommodating heterogeneity and nonlinearity in price effects for predicting brand sales and profits , *European Journal of Operational Research* **246**(1): 232 – 241.
URL: <http://www.sciencedirect.com/science/article/pii/S0377221715001678>
- Lang, S. and Sunder, M. (2003). Non-parametric regression with bayesx: a flexible estimation of trends in human physical stature in 19th century america, *Economics & Human Biology* **1**(1): 77 – 89.
URL: <http://www.sciencedirect.com/science/article/pii/S1570677X02000096>
- Lang, S., Umlauf, N., Wechselberger, P., Harttgen, K. and Kneib, T. (2014). Multilevel structured additive regression, *Statistics and Computing* **24**(2): 223–238.
URL: <http://dx.doi.org/10.1007/s11222-012-9366-0>
- Liu, L., Oza, S., Hogan, D., Chu, Y., Perin, J., Zhu, J., Lawn, J. E., Cousens, S., Mathers, C. and Black, R. E. (2016). Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals, *The Lancet* **388**(10063): 3027 – 3035.
URL: <http://www.sciencedirect.com/science/article/pii/S0140673616315938>
- Lozano, R., Wang, H., Foreman, K. J., Rajaratnam, J. K., Naghavi, M., Marcus, J. R., Dwyer-Lindgren, L., Lofgren, K. T., Phillips, D., Atkinson, C., Lopez, A. D. and Murray, C. J. (2011). Progress towards Millennium Development Goals 4 and 5 on maternal and child mortality: an updated systematic analysis, *The Lancet* **378**(9797): 1139 – 1165.
URL: <http://www.sciencedirect.com/science/article/pii/S0140673611613378>
- Misselhorn, M. and Harttgen, K. (2006). A Multilevel Approach to Explain Child Mortality and Undernutrition in South Asia and Sub-Saharan Africa, *Technical Report 20*, Ibero America Institute

- for Economic Research.
URL: <http://hdl.handle.net/10419/19847>
- Mosley, W. H. and Chen, L. C. (1984). An analytical framework for the study of child survival in developing countries, *Population and Development Review* **10**: 25–45.
URL: <http://www.jstor.org/stable/2807954>
- Pelletier, D. L. and Frongillo, E. A. (2003). Changes in Child Survival Are Strongly Associated with Changes in Malnutrition in Developing Countries, *The Journal of Nutrition* **133**(1): 107–119.
URL: <http://jn.nutrition.org/content/133/1/107.abstract>
- Rutstein, S. (2005). Effects of preceding birth intervals on neonatal, infant and under-five years mortality and nutritional status in developing countries: evidence from the demographic and health surveys, *International Journal of Gynecology & Obstetrics* **89**: S7 – S24. Optimal Birth Spacing: New Evidence on Newborn, Infant, Child and Maternal Health.
URL: <http://www.sciencedirect.com/science/article/pii/S0020729204004710>
- Rutstein, S. O. (2000). Factors associated with trends in infant and child mortality in developing countries during the 1990s, *Bulletin of the World Health Organization* **78**: 1256 – 1270.
URL: http://www.scielo.org/scielo.php?script=sci_arttext&pid=S0042-9686200001000011&nrm=iso
- Rutstein, S. O. and Rojas, G. (2006). Guide to DHS Statistics: Demographic and Health Surveys Methodology, *Technical report*, ORC Macro: Calverton, Maryland.
URL: http://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_29Oct2012_DHSG1.pdf
- Sahn, D. E. and Stifel, D. (2003). Exploring alternative measures of welfare in the absence of expenditure data, *Review of Income and Wealth* **49**(4): 463–489.
URL: <http://dx.doi.org/10.1111/j.0034-6586.2003.00100.x>
- Schultz, T. P. (1984). Studying the Impact of Household Economic and Community Variables on Child Mortality, *Population and Development Review* **10**: 215–235.
URL: <http://www.jstor.org/stable/2807962>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit, *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **64**(4): 583–639.
URL: <http://www.jstor.org/stable/3088806>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2014). The deviance information criterion: 12 years on, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **76**(3): 485–493.
URL: <http://dx.doi.org/10.1111/rssb.12062>
- Umlauf, N., Adler, D., Kneib, T., Lang, S. and Zeileis, A. (2015). Structured Additive Regression Models: An R Interface to BayesX, *Journal of Statistical Software* **63**(1): 1–46.
URL: <https://www.jstatsoft.org/index.php/jss/article/view/v063i21>
- Vollmer, S., Harttgen, K., Kupka, R. and Subramanian, S. V. (2017). Levels and trends of childhood undernutrition by wealth and education according to a Composite Index of Anthropometric Failure: evidence from 146 Demographic and Health Surveys from 39 countries, *BMJ Global Health* **2**(2).
URL: <http://gh.bmj.com/content/2/2/e000206>
- Wang, L. (2003). Determinants of child mortality in LDCs: Empirical findings from demographic and health surveys, *Health Policy* **65**(3): 277 – 299.
URL: <http://www.sciencedirect.com/science/article/pii/S0168851003000393>
- Wechselberger, P., Lang, S. and Steiner, W. J. (2008). Additive Models with Random Scaling factors: Applications to Modeling Price Response Functions, *Austrian Journal of Statistics* **37**(3 & 4): 255–270.
URL: https://www.researchgate.net/profile/Stefan_Lang5/publication/5127479_Additive_models_with_random_scaling_factors_applications_to_modeling_price_response_functions/links/00b7d51d522c03df99000000.pdf
- Wolpin, K. I. (1997). *Determinants and consequences of the mortality and health of infants and children*, Vol. 1 A, Elsevier, Amsterdam, North Holland.
URL: <http://www.sciencedirect.com/science/article/pii/S1574003X9780027X>

- You, D., Hug, L., Ejdemyr, S., Idele, P., Hogan, D., Mathers, C., Gerland, P., New, J. and Alkema, L. (2015). Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation, *The Lancet* **386**(10010): 2275 – 2286.
URL: <http://www.sciencedirect.com/science/article/pii/S0140673615001208>
- You, D., Jones, G., Hill, K., Wardlaw, T. and Chopra, M. (2010). Levels and trends in child mortality, 1990-2009, *The Lancet* **376**(9745): 931–933.
URL: [http://dx.doi.org/10.1016/S0140-6736\(10\)61429-8](http://dx.doi.org/10.1016/S0140-6736(10)61429-8)
- You, D., New, J. and Wardlaw, T. (2012). Levels and trends in child mortality. Report 2012. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation., *Technical report*, United Nations.
- You, D., New, J. and Wardlaw, T. (2014). Levels and trends in child mortality. Report 2014. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation., *Technical report*, United Nations.
- You, D., New, J. and Wardlaw, T. (2015). Levels and trends in child mortality. Report 2015. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation., *Technical report*, United Nations.

6. Figures and Tables

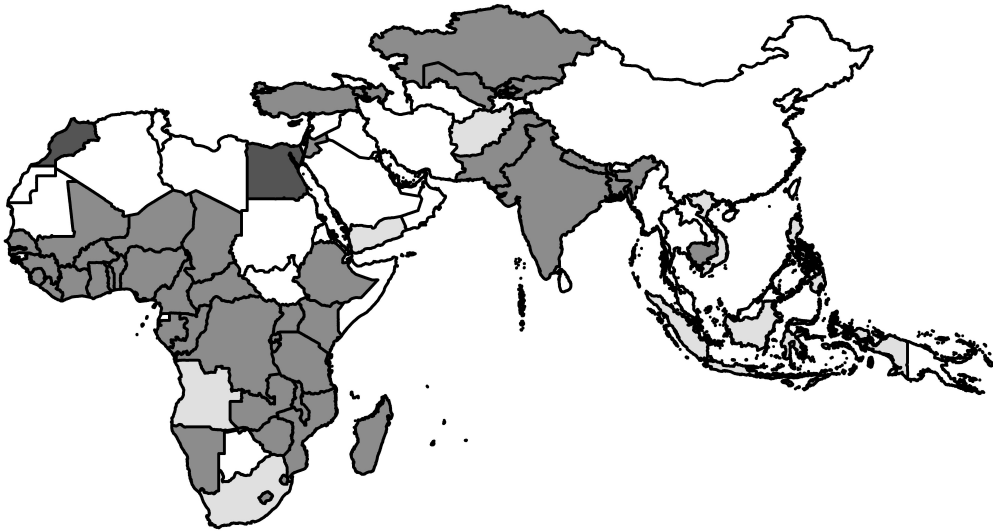


Fig. 1: Map of included countries: Boundary information of the map is taken from www.gadm.org; Gray colored countries indicate included countries (48); Dark gray colored countries indicate Arabian countries not located (completely) in Asia (2); Light gray countries indicate countries with missing variables or countries where regional structure changed considerably and a harmonization was not possible (8).

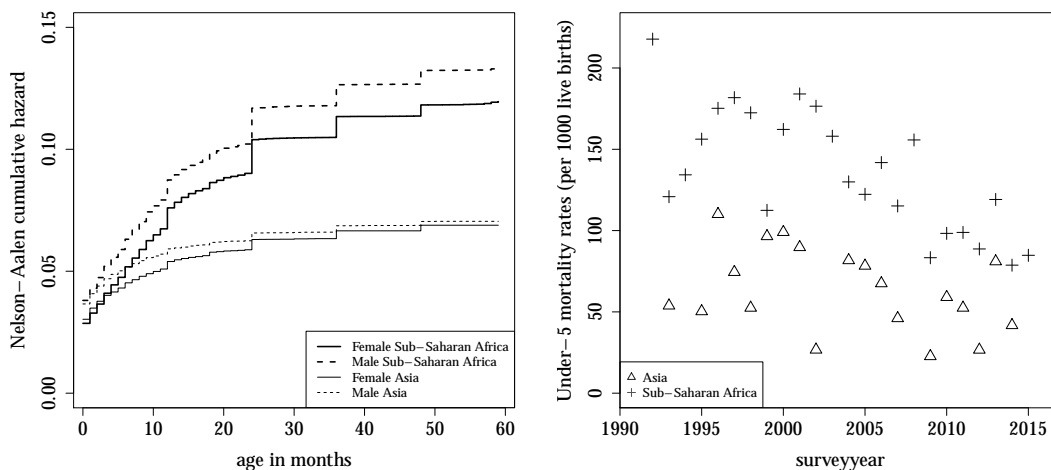


Fig. 2: Nelson-Aalen cumulative hazard by sex and continent, and development under-5 mortality over time. **Source:** DHS; calculations by authors.

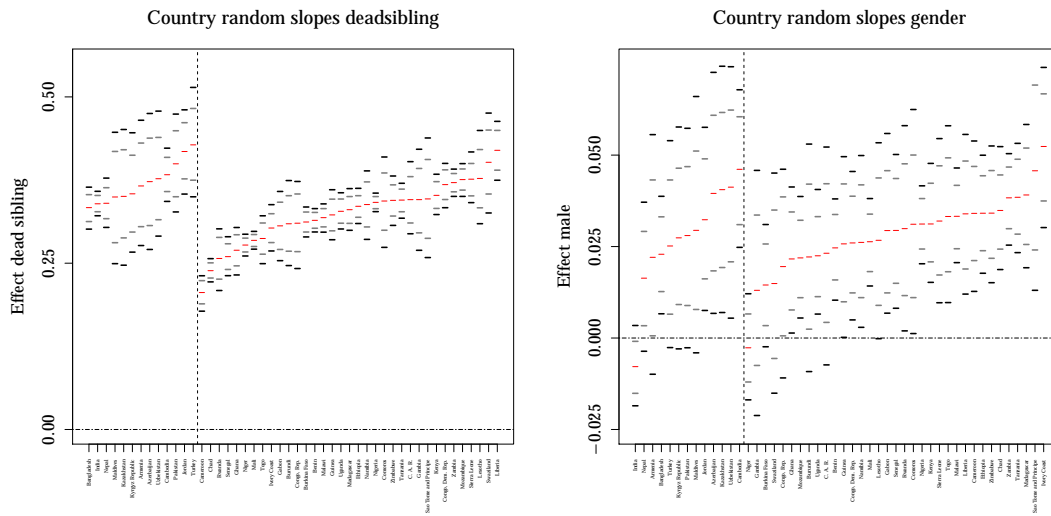


Fig. 3: Categorical covariates, country random slopes: Figure depicts the posterior means of the categorial variables *deadsib*, and *male* together with 80% and 95% simultaneous credible intervals. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

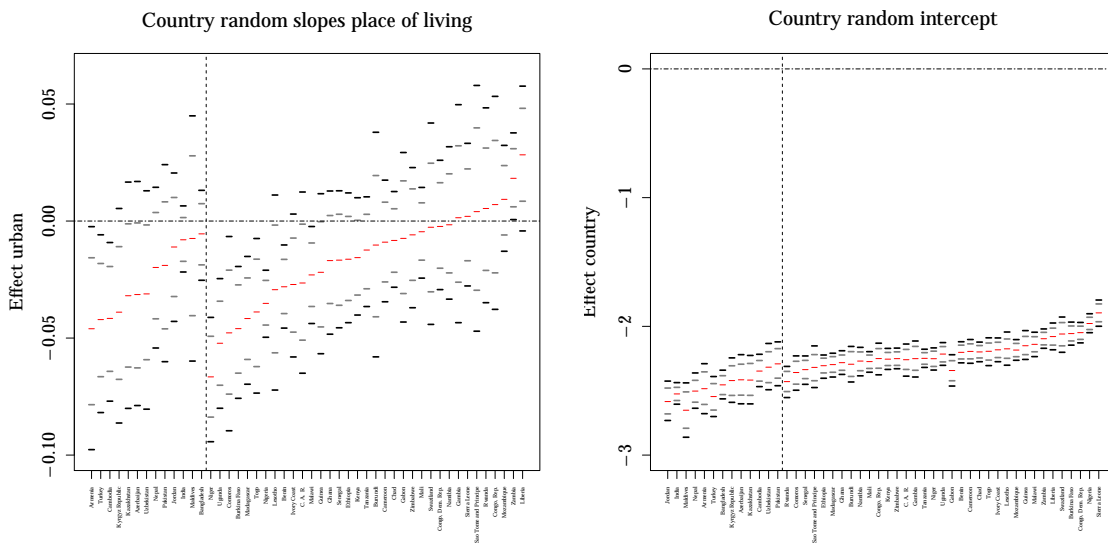


Fig. 4: Categorical covariates, country random slopes and random intercepts: Figure depicts the posterior means of the categorial variables *male*, and the random intercept together with 80% and 95% simultaneous credible intervals. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

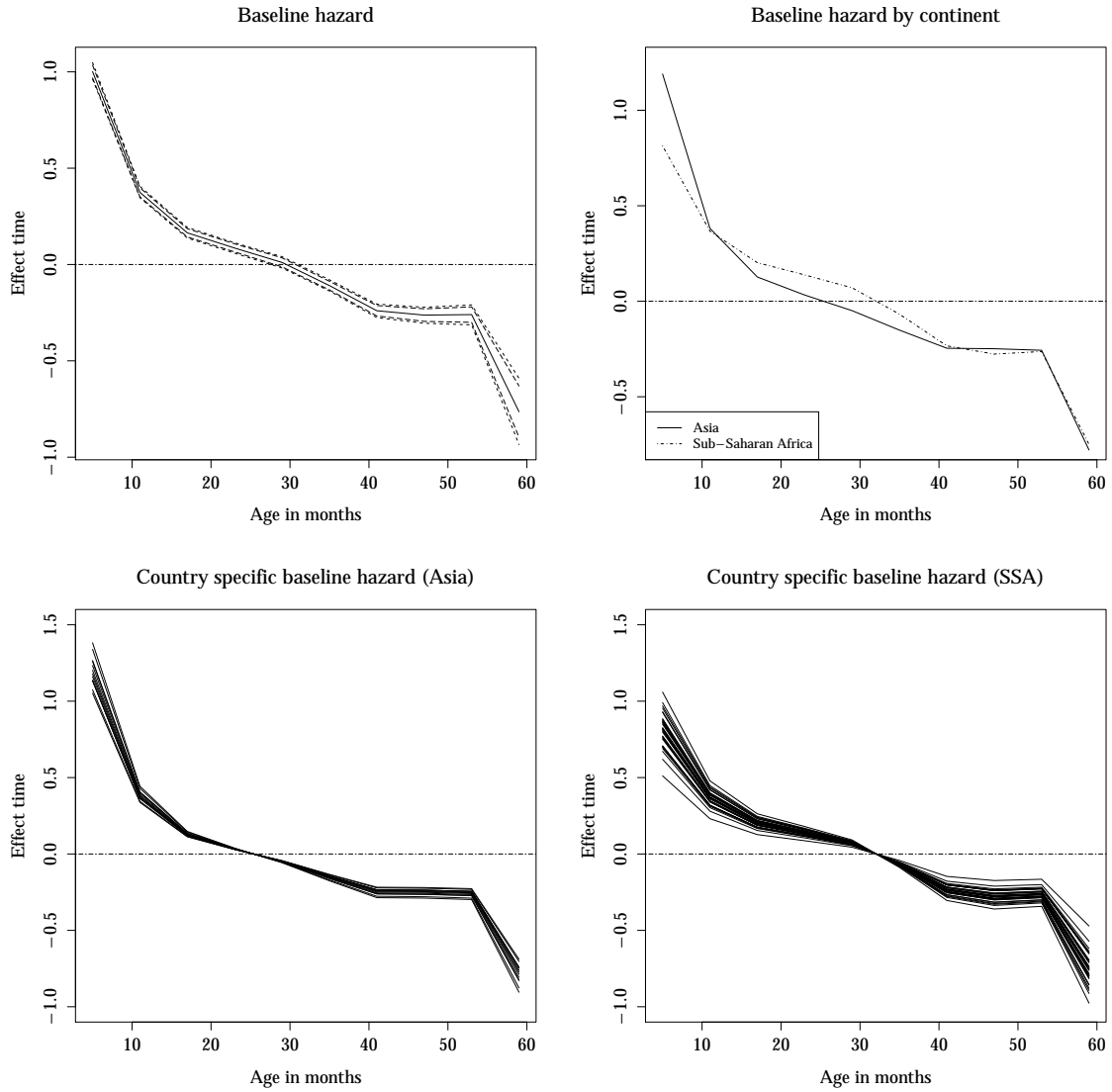


Fig. 5: Non-linear effect of the baseline hazard at level-1. Shown is the mean effect, together with 95% and 80% simultaneous credible intervals. Additionally, the continent specific mean effects, and the with the multiplicative country random effects scaled effects for Asia and Sub-Saharan Africa are shown. **Source:** DHS; calculations by the authors.

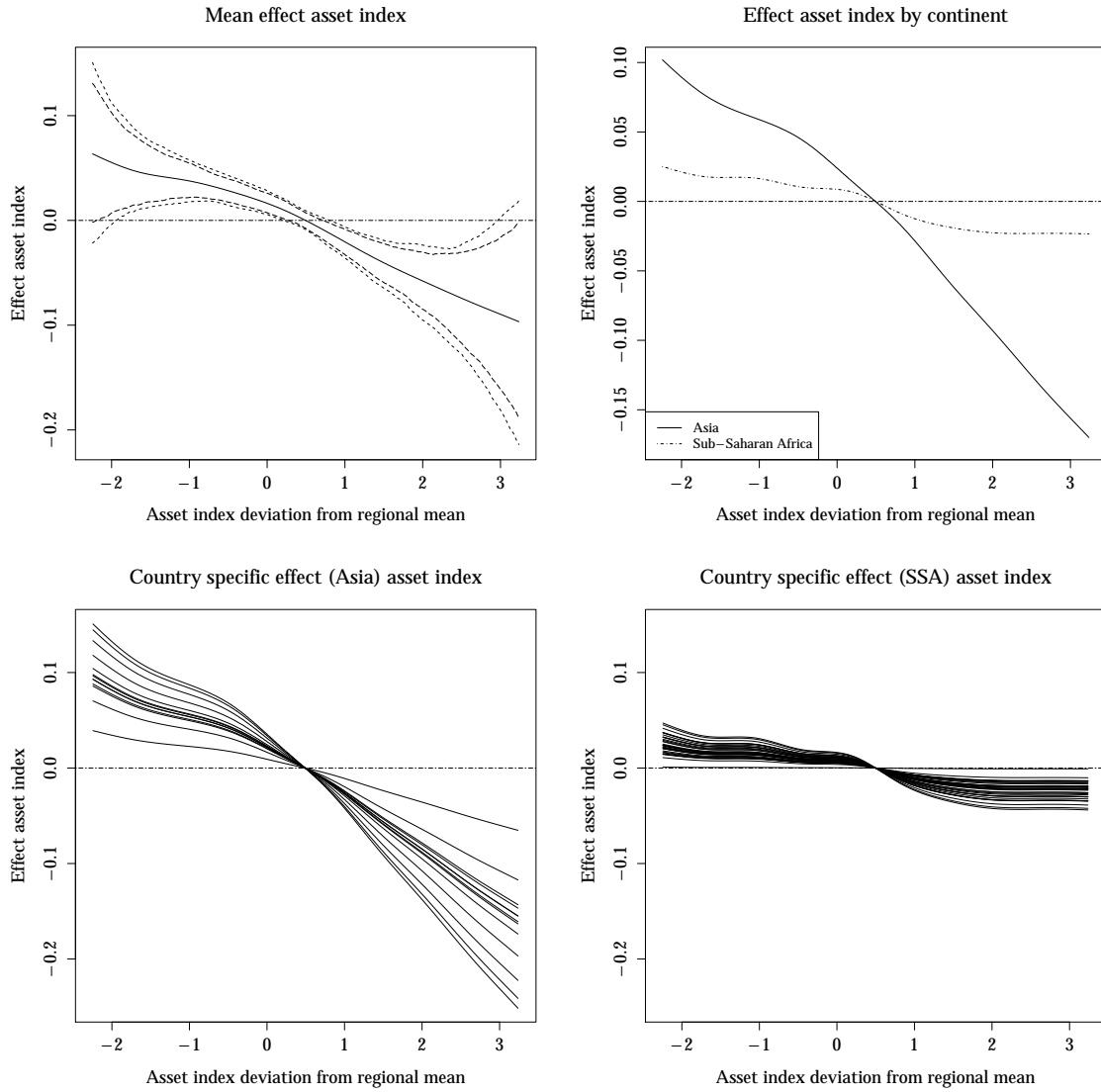


Fig. 6: Non-linear effect of the asset index at level-1. Shown is the mean effect, together with 95% and 80% simultaneous credible intervals. Additionally, the continent specific mean effects, and the with the multiplicative country random effects scaled effects for Asia and Sub-Saharan Africa are shown. **Source:** DHS; calculations by the authors.

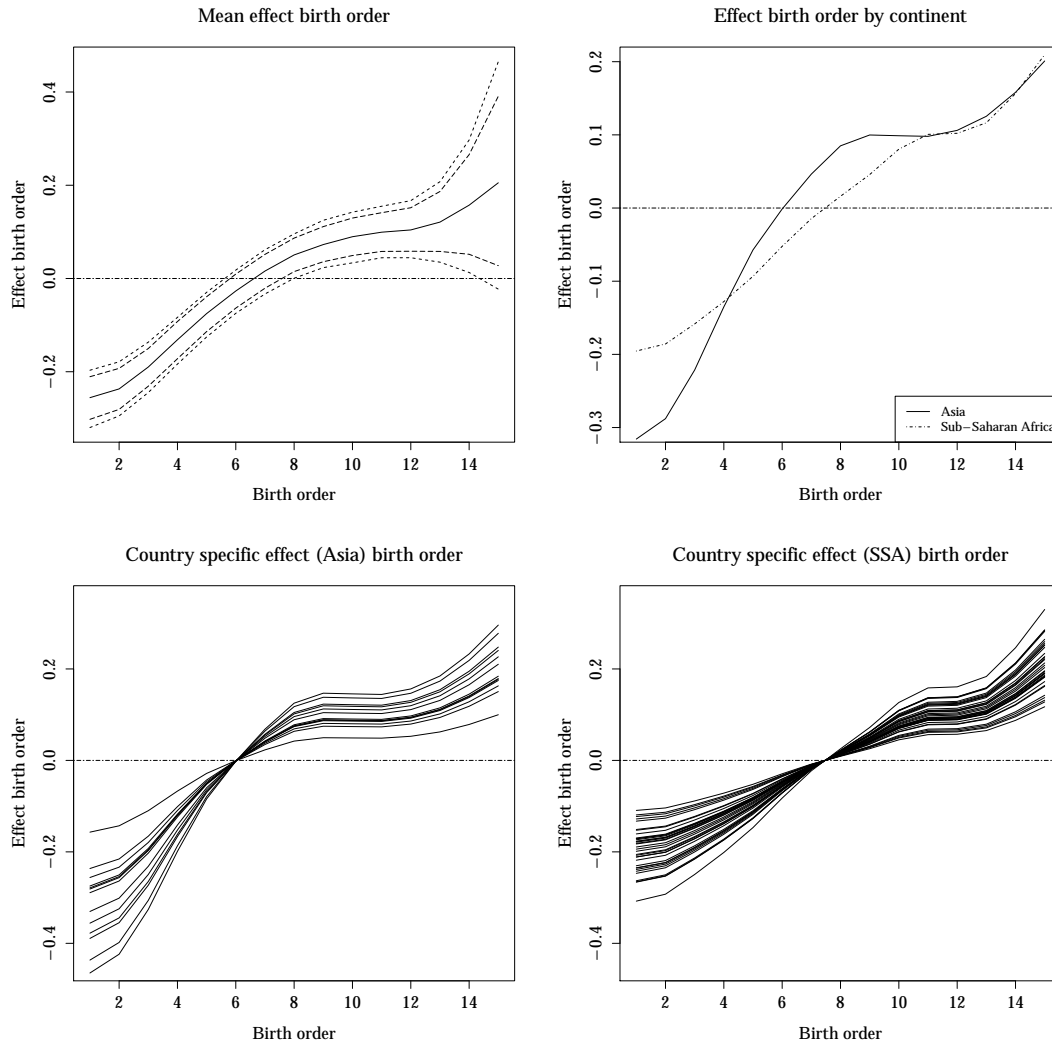


Fig. 7: Non-linear effect of the birth order within the household at level-1. Shown is the mean effect together with 95% and 80% simultaneous credible intervals. Additionally, the continent specific mean effects, and the with the multiplicative country random effects scaled effects for Asia and Sub-Saharan Africa are shown. **Source:** DHS; calculations by the authors.

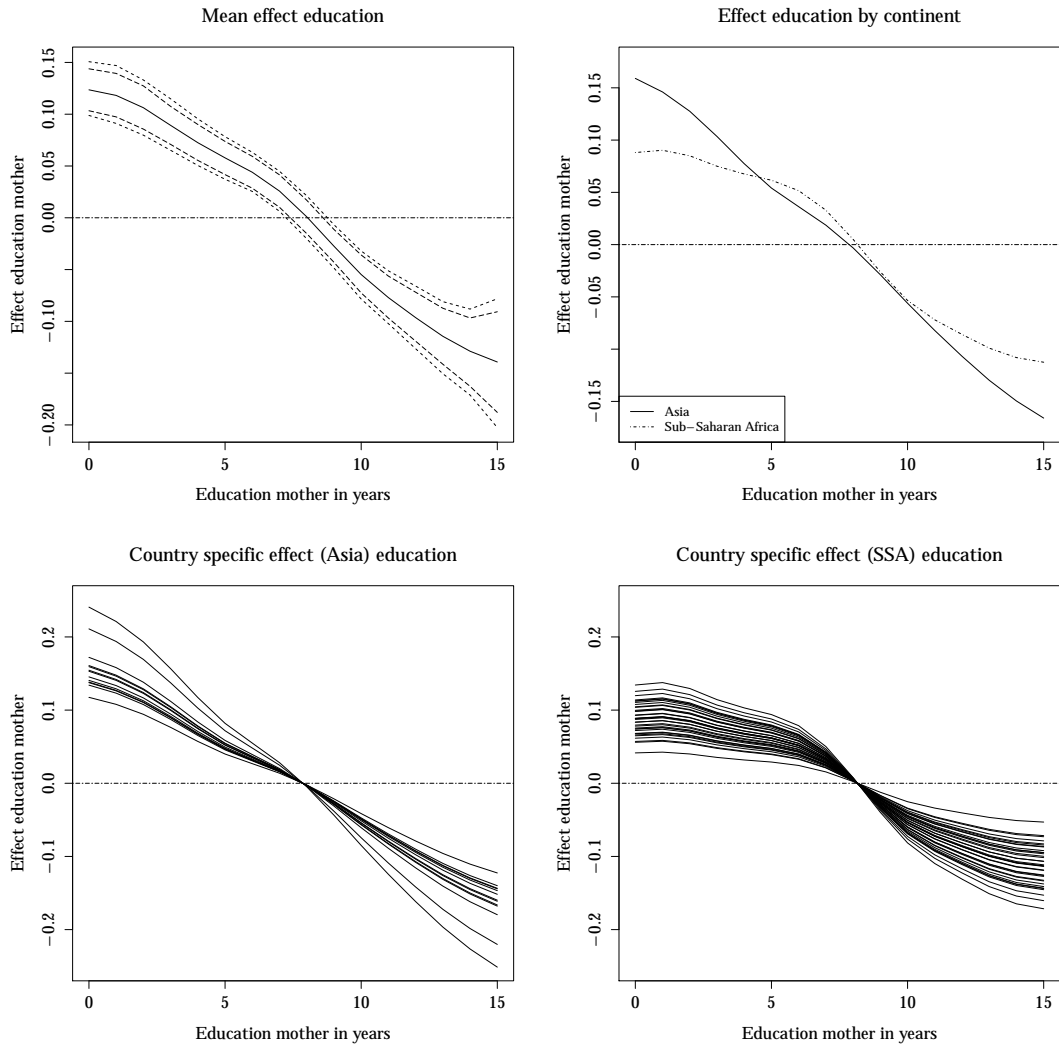


Fig. 8: Non-linear effect of the mother's years of education at level-1. Shown is the mean effect, together with 95% and 80% simultaneous credible intervals. Additionally, the continent specific mean effects, and the with the multiplicative country random effects scaled effects for Asia and Sub-Saharan Africa are shown. **Source:** DHS; calculations by the authors.

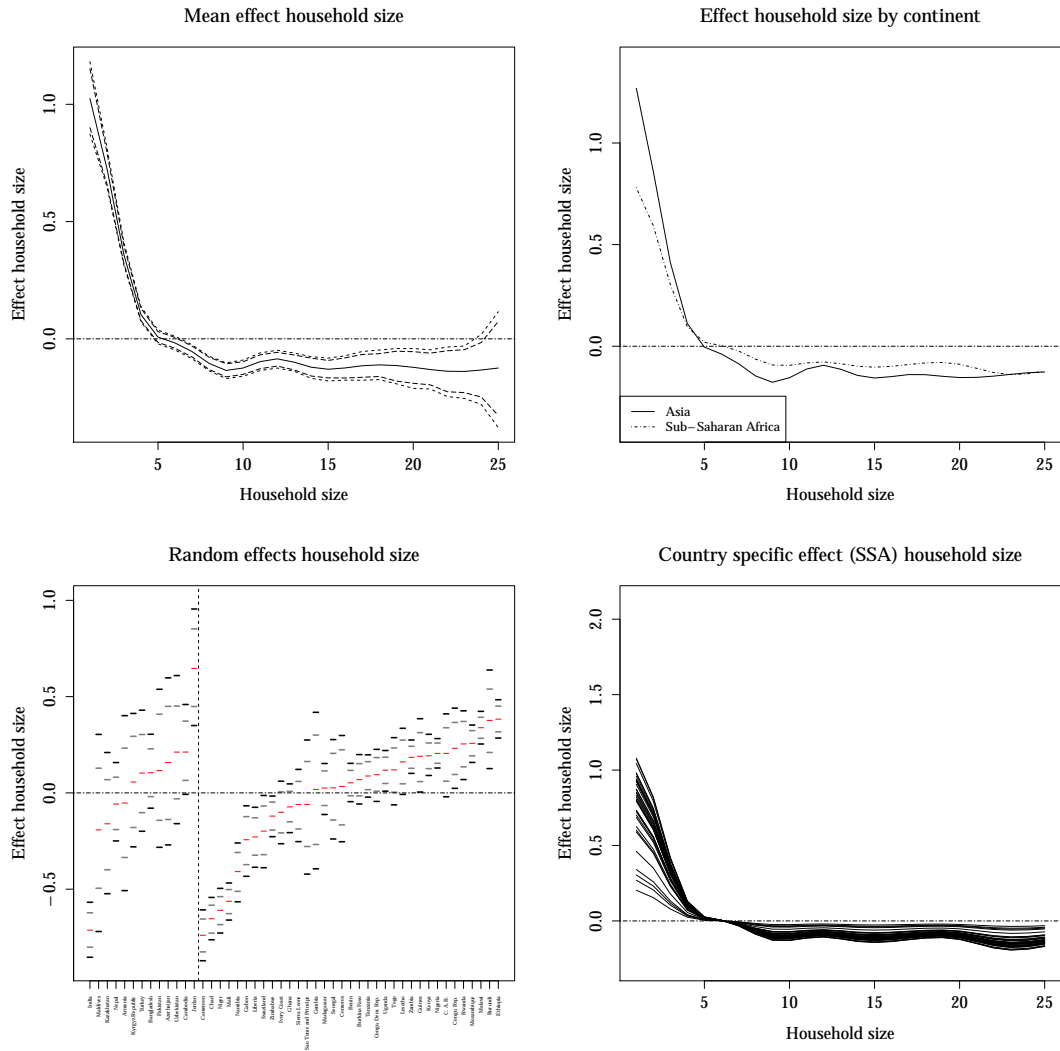


Fig. 9: Non-linear effect of the size of the household at level-1. Shown is the mean effect, together with 95% and 80% simultaneous credible intervals. Additionally, the continent specific mean effects, and the with the multiplicative country random effects scaled effects for Asia and Sub-Saharan Africa are shown. **Source:** DHS; calculations by the authors.

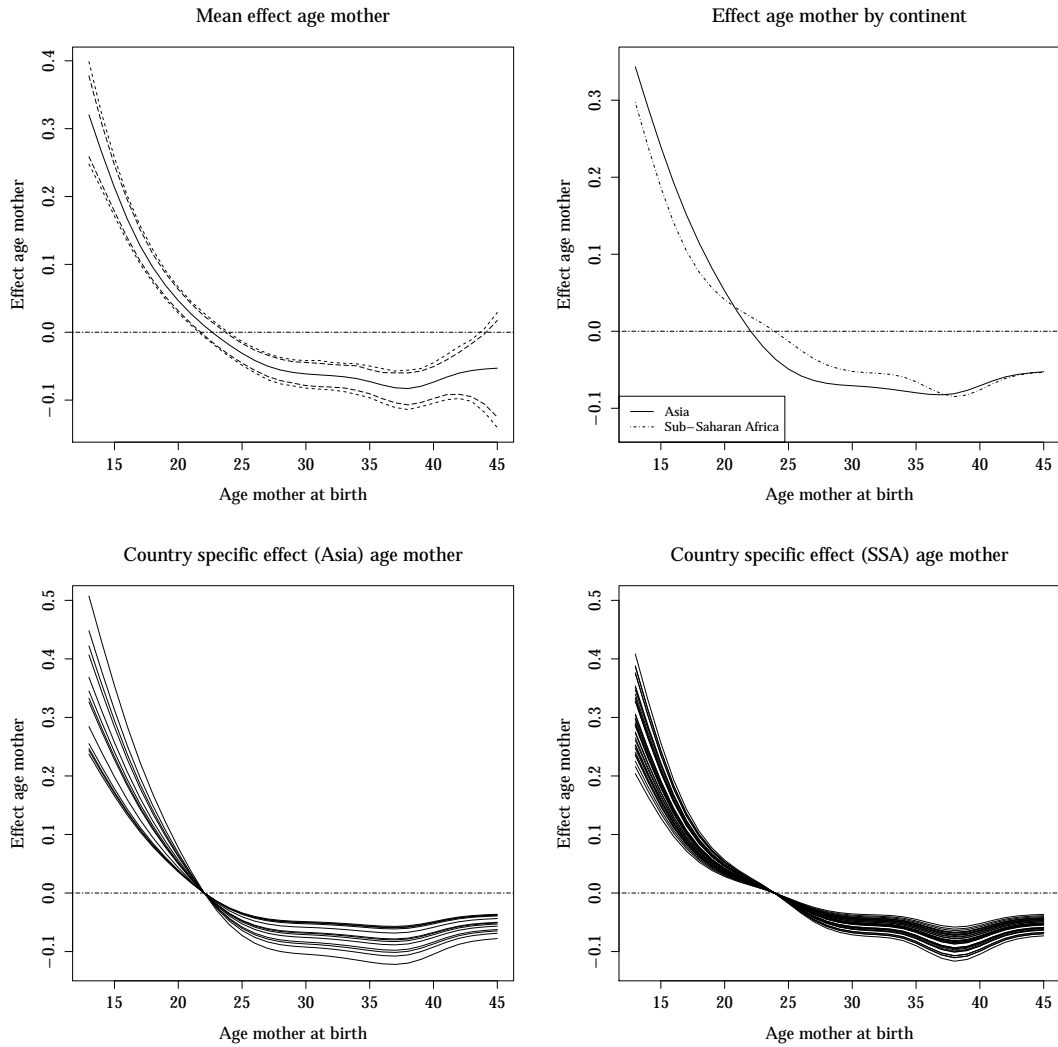


Fig. 10: Non-linear effect of the age of the mother at birth at level-1. Shown is the mean effect, together with 95% and 80% simultaneous credible intervals. Additionally, the continent specific mean effects, and the with the multiplicative country random effects scaled effects for Asia and Sub-Saharan Africa are shown. **Source:** DHS; calculations by the authors.

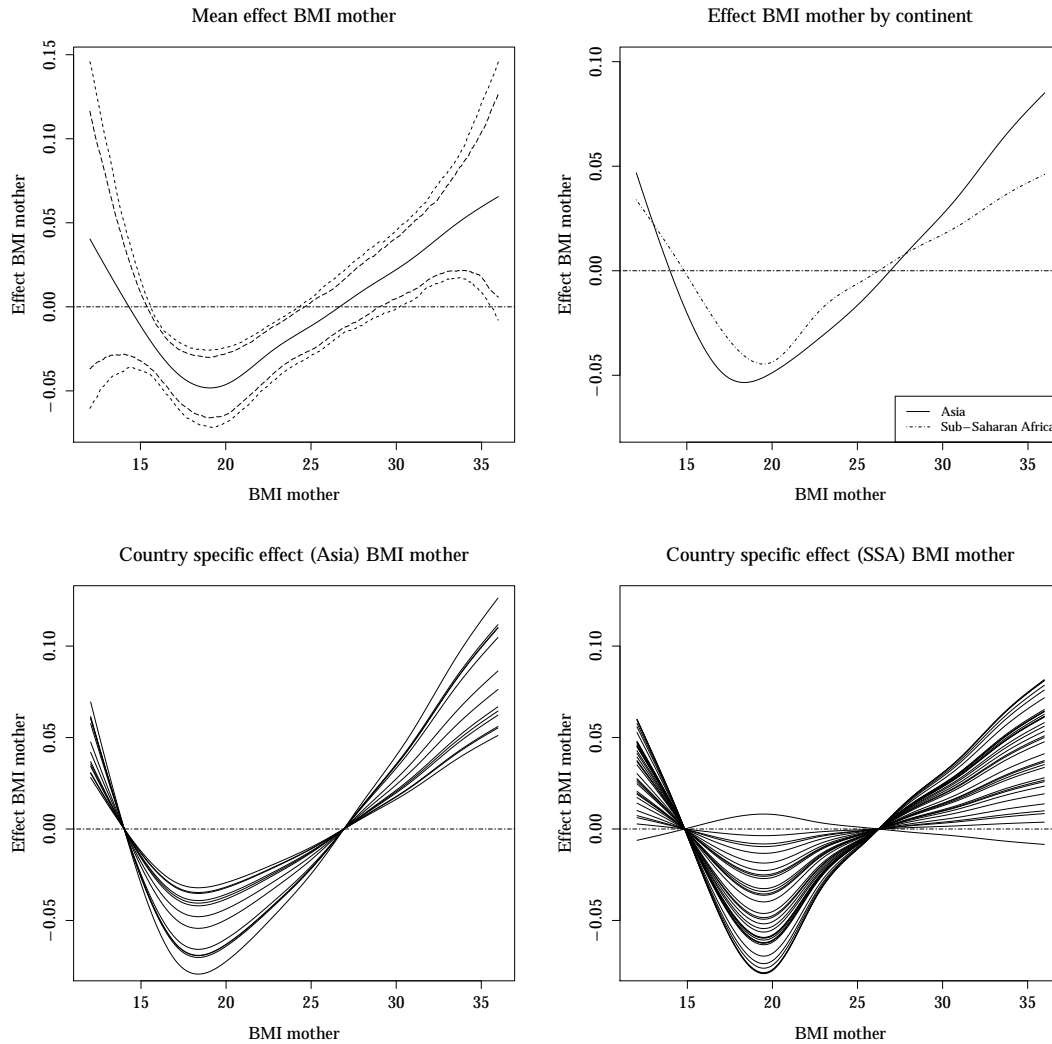


Fig. 11: Non-linear effect of the mother's BMI at level-1. Shown is the mean effect, together with 95% and 80% simultaneous credible intervals. Additionally, the continent specific mean effects, and the with the multiplicative country random effects scaled effects for Asia and Sub-Saharan Africa are shown. **Source:** DHS; calculations by the authors.

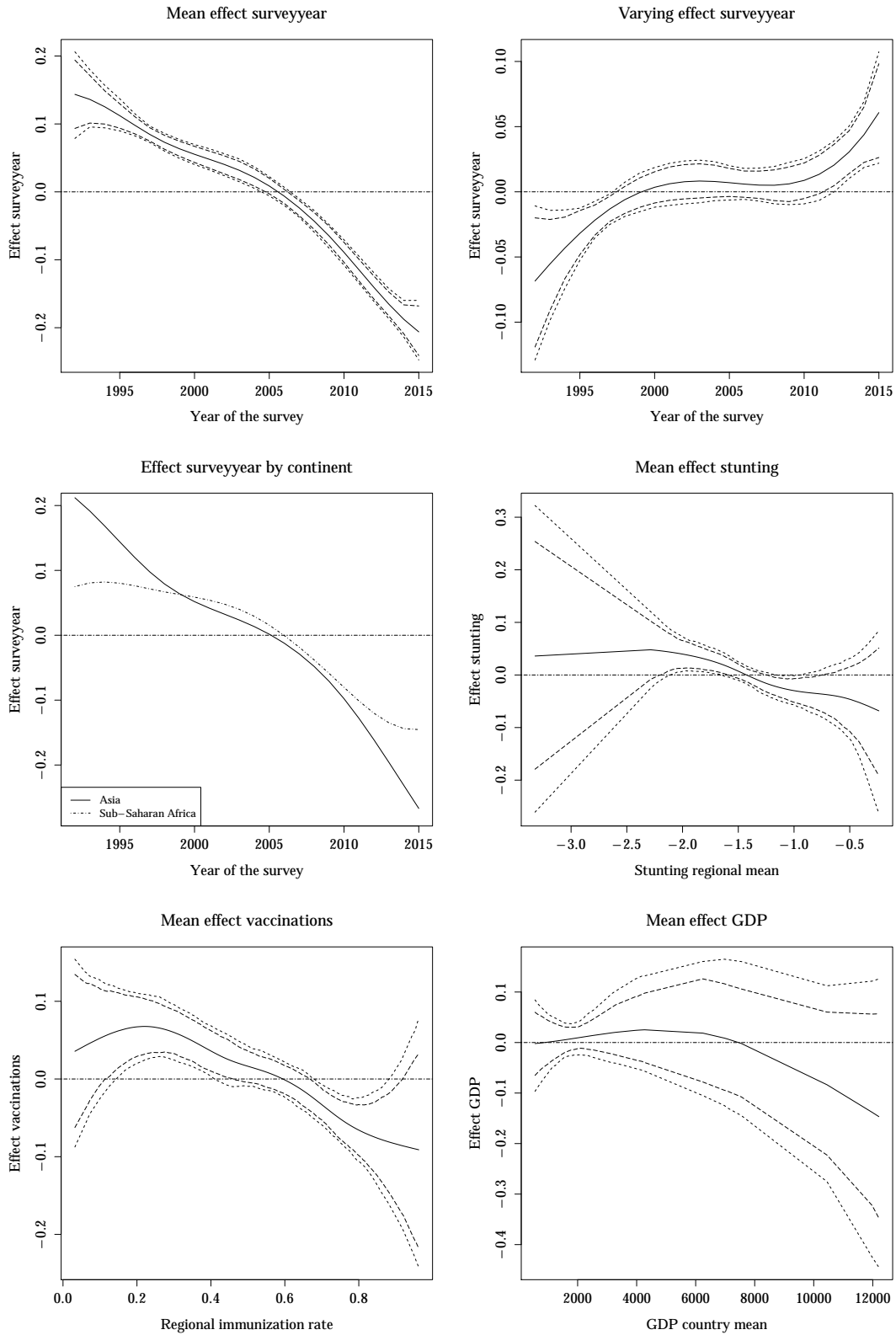


Fig. 12: Non-linear effects of the survey year at level-1, the nutritional status, and regional immunization rate at level-2, and GDP at level-3. Shown are the mean effects, and for the surveyyear the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect of the geographic regions. **Source:** DHS, and Penn World Table 9.0; calculations by the authors.

Tab. 1: Under-5 mortality rates by continent

| Sample | Number of observations (n) | Mortality rate (per 1,000 children) |
|--------------------|----------------------------|-------------------------------------|
| Total | 824,017 | 115 |
| Asia | 220,532 | 72 |
| Sub-Saharan Africa | 603,485 | 131 |

The sample contains 824,017 observations nested in 358 regions and 48 countries. **Source:** DHS; calculations by authors.

Tab. 2: Expected effects of the covariates on under-5 mortality

| Definition variable | Type | Level | Expected effect |
|-------------------------------------|----------------|---------|-----------------|
| Child has dead siblings | discrete (d) | Level-1 | positive |
| Household has female head | d | Level-1 | positive |
| Sex of child | d | Level-1 | positive |
| Place of living | d | Level-1 | negative |
| Asset index deviation regional mean | continuous (d) | Level-1 | decreasing |
| Birth order within household | c | Level-1 | U-shaped |
| Years of education mother | c | Level-1 | decreasing |
| Number of people in household | c | Level-1 | increasing |
| Age of mother at birth | c | Level-1 | U-shaped |
| BMI mother | c | Level-1 | U-shaped |
| Preceding birth interval | c | Level-1 | U-shaped |
| Year of the survey | c | Level-1 | decreasing |
| Z-stunting regional mean | c | Level-2 | decreasing |
| Regional immunization coverage | c | Level-2 | decreasing |
| Continent | d | Level-3 | positive |
| GDP p.c. country mean | c | Level-3 | decreasing |

Categorical covariates: The expected effect of the discrete variables on child mortality are assumed to be positive, negative or neutral. **Continuous covariates:** The expected effect of the continuous variables on child mortality are assumed to be increasing, decreasing or U-shaped.

Tab. 3: Descriptive statistic covariates variables

| Definition covariates | Type | Mean, % | std. dev. | <i>n</i> at Level | Level |
|---|------|----------|-----------|-------------------|---------|
| Child has older dead sibling (=1) | d | 10.2% | | 824,017 | Level-1 |
| Sex of child (1=male) | d | 50.78% | | 824,017 | Level-1 |
| Place of living (1=urban) | d | 28.74% | | 824,017 | Level-1 |
| Continent (1=Sub-Saharan Africa) | d | 73.24% | | 48 | Level-3 |
| Asset index deviation from regional mean | c | -0.02 | 0.79 | 824,017 | Level-1 |
| Birth order of child | c | 3.51 | 2.4 | 824,017 | Level-1 |
| Years of education mother | c | 3.9 | 4.23 | 824,017 | Level-1 |
| Number of household members | c | 6.86 | 3.49 | 824,017 | Level-1 |
| BMI mother | c | 21.87 | 3.6 | 824,017 | Level-1 |
| Age mother at birth (in years) | c | 25.35 | 6.6 | 824,017 | Level-1 |
| Preceding birth interval mother | c | 37.95 | 20.32 | 627,201 | Level-1 |
| Z-score stunting regional (reg.) mean | c | -1.4 | 0.42 | 358 | Level-2 |
| Reg. vaccination rate of specified vaccines | c | 0.55 | 0.21 | 358 | Level-2 |
| Real per capita GDP | c | 2,923.15 | 2,841.73 | 48 | Level-3 |

The sample contains 824,017 observations in 358 regions and 48 countries. **Source:** DHS, Penn World Table 9.0, World Bank WDI; calculations by authors.

Tab. 4: Estimated results of the categorial covariates

| Definition Variable | Mean | Std. Dev. | 95 _{lb} % CI | 95 _{ub} % CI |
|----------------------------------|--------|-----------|-----------------------|-----------------------|
| Overall intercept | -2.337 | 0.033 | -2.403 | -2.273 |
| Continent (1=SSA) | 0.128 | 0.032 | 0.067 | 0.191 |
| Older sibling of child died (=1) | 0.337 | 0.009 | 0.319 | 0.354 |
| Gender (male=1) | 0.028 | 0.004 | 0.021 | 0.035 |
| Urban area of living (=1) | -0.018 | 0.005 | -0.029 | -0.009 |

Estimation results of categorial covariates. All covariates are effect coded. **Source:** DHS; calculation by authors.

Tab. 5: Results model with interaction place of living | continent

| Definition Variable | mean | std. dev. | 95 _{lb} % CI | 95 _{ub} % CI |
|---------------------|--------|-----------|-----------------------|-----------------------|
| Urban SSA | 0.078 | 0.053 | -0.023 | 0.163 |
| Rural SSA | 0.111 | 0.053 | 0.01 | 0.194 |
| Urban Asia | -0.11 | 0.053 | -0.197 | -0.007 |
| Rural Asia | -0.079 | | | |

Results categorial covariates interaction place of living, continent. All covariates are effect coded. **Source:** DHS; calculation by authors.

A. Under-5 mortality rates by country and survey year

Tab. A.1: Number of observations (n) and under-5 mortality rates per 1,000 live births (U5M) for the five-year period preceding the survey by country, and survey year.

| Country | Year | n | U5M | Country | Year | n | U5M | Country | Year | n | U5M |
|------------------|------|--------|-----|-----------------|------|--------|-----|-----------------------|------|--------|-----|
| Armenia | 2000 | 1,493 | 41 | Guinea | 2012 | 3,644 | 124 | Nepal | 1996 | 5,966 | 133 |
| Armenia | 2005 | 1,259 | 23 | India | 1999 | 48,935 | 96 | Nepal | 2001 | 6,864 | 89 |
| Azerbaijan | 2006 | 2,035 | 49 | India | 2006 | 46,043 | 69 | Nepal | 2006 | 5,650 | 64 |
| Bangladesh | 1997 | 5,764 | 114 | Ivory Coast | 1994 | 5,273 | 158 | Niger | 1992 | 6,538 | 289 |
| Bangladesh | 2000 | 6,612 | 85 | Ivory Coast | 1999 | 501 | 131 | Niger | 1998 | 6,679 | 265 |
| Bangladesh | 2004 | 6,768 | 80 | Ivory Coast | 2012 | 3,748 | 117 | Niger | 2006 | 4,311 | 164 |
| Bangladesh | 2007 | 5,975 | 67 | Jordan | 1997 | 5,377 | 34 | Niger | 2012 | 5,678 | 117 |
| Bangladesh | 2011 | 8,442 | 53 | Jordan | 2002 | 4,735 | 27 | Nigeria | 2003 | 5,784 | 205 |
| Bangladesh | 2014 | 7,665 | 47 | Jordan | 2007 | 4,020 | 18 | Nigeria | 2008 | 26,460 | 165 |
| Benin | 1996 | 4,092 | 170 | Jordan | 2009 | 3,324 | 28 | Nigeria | 2013 | 29,614 | 127 |
| Benin | 2001 | 5,132 | 157 | Jordan | 2012 | 4,721 | 25 | Pakistan | 2013 | 3,408 | 81 |
| Benin | 2006 | 15,024 | 123 | Kazakhstan | 1995 | 1,333 | 52 | Rwanda | 2010 | 4,451 | 79 |
| Benin | 2012 | 12,879 | 79 | Kazakhstan | 1999 | 597 | 63 | Rwanda | 2015 | 3,730 | 50 |
| Burkina Faso | 2003 | 10,115 | 179 | Kenya | 1993 | 5,635 | 98 | Sao Tome and Principe | 2009 | 1,695 | 59 |
| Burkina Faso | 2010 | 7,528 | 129 | Kenya | 1998 | 5,259 | 112 | Senegal | 1993 | 4,994 | 132 |
| Burundi | 2010 | 3,523 | 91 | Kenya | 2003 | 5,219 | 112 | Senegal | 2005 | 3,022 | 111 |
| C. A. R. | 1995 | 3,840 | 149 | Kenya | 2009 | 5,816 | 77 | Senegal | 2011 | 3,964 | 78 |
| Cambodia | 2000 | 4,038 | 142 | Kenya | 2014 | 9,738 | 53 | Sierra Leone | 2008 | 2,692 | 145 |
| Cambodia | 2005 | 4,088 | 94 | Kyrgyz Republic | 1997 | 1,886 | 64 | Sierra Leone | 2013 | 5,746 | 154 |
| Cambodia | 2010 | 4,001 | 60 | Kyrgyz Republic | 2012 | 3,400 | 28 | Swaziland | 2007 | 2,537 | 122 |
| Cambodia | 2014 | 4,572 | 34 | Lesotho | 2004 | 1,745 | 120 | Tanzania | 1996 | 6,185 | 98 |
| Cameroon | 1998 | 3,000 | 168 | Lesotho | 2009 | 1,957 | 122 | Tanzania | 2005 | 8,125 | 112 |
| Cameroon | 2004 | 3,942 | 149 | Lesotho | 2014 | 1,532 | 96 | Tanzania | 2010 | 7,718 | 77 |
| Cameroon | 2011 | 5,809 | 130 | Liberia | 2007 | 5,382 | 112 | Tanzania | 2015 | 9,782 | 66 |
| Chad | 1997 | 7,009 | 200 | Liberia | 2013 | 3,849 | 114 | Togo | 1998 | 5,650 | 166 |
| Chad | 2004 | 5,206 | 174 | Madagascar | 1997 | 5,082 | 166 | Togo | 2014 | 3,405 | 92 |
| Chad | 2015 | 11,327 | 136 | Madagascar | 2004 | 5,192 | 82 | Turkey | 1993 | 3,361 | 55 |
| Comoros | 1996 | 1,555 | 107 | Madagascar | 2009 | 6,045 | 84 | Turkey | 1998 | 2,990 | 56 |
| Comoros | 2012 | 2,753 | 51 | Malawi | 1992 | 4,241 | 233 | Uganda | 1995 | 6,325 | 160 |
| Congo, Dem. Rep. | 2007 | 4,336 | 151 | Malawi | 2000 | 11,454 | 186 | Uganda | 2001 | 6,350 | 149 |
| Congo, Dem. Rep. | 2014 | 9,342 | 121 | Malawi | 2004 | 10,282 | 130 | Uganda | 2006 | 2,844 | 118 |
| Congo, Rep. | 2012 | 4,789 | 68 | Malawi | 2010 | 6,466 | 108 | Uganda | 2011 | 2,414 | 101 |
| Ethiopia | 2000 | 10,510 | 168 | Malawi | 2015 | 5,526 | 60 | Uzbekistan | 1996 | 2,219 | 62 |
| Ethiopia | 2005 | 4,793 | 126 | Maldives | 2009 | 2,991 | 18 | Zambia | 1992 | 6,038 | 190 |
| Ethiopia | 2011 | 11,118 | 99 | Mali | 1996 | 7,176 | 234 | Zambia | 1996 | 6,942 | 196 |
| Gabon | 2000 | 4,004 | 75 | Mali | 2001 | 12,327 | 213 | Zambia | 2002 | 6,690 | 177 |
| Gabon | 2012 | 3,731 | 81 | Mali | 2006 | 13,576 | 178 | Zambia | 2007 | 6,282 | 116 |
| Gambia | 2013 | 3,294 | 49 | Mali | 2013 | 5,254 | 110 | Zambia | 2014 | 13,048 | 68 |
| Ghana | 1993 | 2,972 | 131 | Mozambique | 1997 | 5,538 | 176 | Zimbabwe | 1994 | 3,112 | 89 |
| Ghana | 1998 | 3,148 | 104 | Mozambique | 2003 | 9,452 | 149 | Zimbabwe | 1999 | 3,385 | 111 |
| Ghana | 2003 | 3,590 | 115 | Mozambique | 2011 | 10,856 | 96 | Zimbabwe | 2006 | 5,076 | 91 |
| Ghana | 2008 | 2,865 | 86 | Namibia | 1992 | 3,467 | 84 | Zimbabwe | 2011 | 5,176 | 85 |
| Ghana | 2014 | 2,862 | 67 | Namibia | 2007 | 4,677 | 80 | Zimbabwe | 2015 | 5,563 | 65 |
| Guinea | 2005 | 3,238 | 153 | Namibia | 2013 | 2,245 | 55 | | | | |

Source: DHS data sets; calculation by authors.

B. Selection process

Model selection of the final model was accomplished using as selection criteria Bayesian credible bands Krivobokova et al. (2010), to identify significance of covariates, and the DIC Spiegelhalter et al. (2002), to compare the different specifications. In the model building phase, the MCMC sampler was set to 45,000 and a burn-in period of 5,000 iterations was used. Depending on the complexity of the model, estimation in this stage takes approximately between 1,500 and 2,600 minutes on a laptop operating on Windows basis, and Intel i7-5600U 2.60 GHz processor. In the final model we used 110,000 iterations and a burn-in phase of 10,000 iterations for the MCMC sampler. The DIC compromises a trade-off between the fit of the data, and the complexity of the model. The lower the DIC the better the model describes the data.

In a further step, a variable that enters the model linearly is considered to be significant if the 95% credible band does not fully include zero. If the covariate is modeled as a non-linear effect, the variable is called significant if the 95% credible band does not include the zero line over the complete range of the covariate. Simultaneous Bayesian confidence bands are obtained for effect f_i from the effect's posterior distribution which is observed from the data, where the effect f_i is estimated using MCMC simulation techniques (Krivobokova et al., 2010). Intuitively, the wider the credible bands in a certain range of the data, the more inaccurate f_i describes the data due to the scarce information the data provides. The narrower the credible bands the better the information on the curve of f_i .

To discriminate between the estimated models and to select the model that offers the best trade-off between model complexity and model fit the DIC is consulted, since it can be calculated for hierarchical models. Table B.1 shows the differences of the DIC of the estimated models to the 'baseline model'.

Selection of the baseline model Selecting between a solely additive model and multilevel models with either 2 or 3 levels, the DIC strictly favors the multilevel model with 3 levels of Equation 6 compared to the simple additive model. The DIC of the additive model is roughly 2,000 points higher in comparison to the multilevel model with 3 levels, and without random effects. A comparison of the multilevel model with 3 levels instead of 2 levels, and accounting for the fact that individuals are nested into regions, which are clustered into countries reveals the following: Incorporating, besides the country specific structure of the data, the regional specific structure yields an improvement, lowering the DIC by approximately 800.

Allowing for different effect sizes of countries So far cluster specific heterogeneity is only considered through the country, respectively, region specific random intercept. As mentioned earlier, it is possible to allow for cluster specific effect sizes for all covariates. Doing so, and, including country random slopes as for instance, described in Brezger et al. (2005) into Equation 6 shows a clear decrease of the DIC of more than 2,600 in comparison to the models without random slopes. Additionally, modeling country

specific heterogeneity of the continuous covariates through multiplicative random effects yields a further improvement. The DIC decreases by around 1,200 points in comparison to the models capturing heterogeneity with random slopes. This clearly emphasizes the advantage of modeling country specific heterogeneity with multiplicative random effects.

Taking advantage of the rule of thumb for the DIC by Spiegelhalter et al. (2002) the multilevel model that includes covariates on the individual, regional and country level is always favorable in comparison to models where only the the individual level or the individual and country level is considered. Furthermore, accounting for heterogeneity yields to further improvements of the model. Accounting for heterogeneity is most efficient when multiplicative random effects are used.

In favor of this decision rule, results are discussed for the ‘baseline’ model of Equation 6 including multiplicative country random effects, and omitting insignificant terms. Besides this ‘baseline’ model, we estimated several models similar to the afore mentioned model, allowing for interactions between categorical covariates and the geopolitical region (e.g. gender—continent, place of living—continent), and also one model distinguishing between South Asian and Western/Central Asian countries. However, in comparison to the ‘baseline’ model, the results for these specifications are rather similar with respect to shape and magnitude. Additionally, only the model which distinguishes within Asia between South Asia and Western/Central Asia has a more favorable DIC than the ‘baseline’ model. Accordingly, we will present and discuss the results of the ‘baseline’ model and mention important results stemming from these models.

Tab. B.1: Model selection: Differences DIC to baseline model

| Model | Δ DIC |
|--|--------------|
| additive model (AM) | 0 |
| multilevel AM (MAM) with 2 levels | -1,209 |
| MAM Equation 6 | -2,002 |
| MAM with 2 levels, including country random slopes | -3,796 |
| Equation 6, including country random slopes | -4,439 |
| MAM with 2 levels, including multiplicative country random effects (MRE) | -4,954 |
| MAM Equation 6, including MRE | -5,624 |
| MAM Equation 6, including MRE, omitting insignificant varying coefficient terms | -5,632 |
| MAM Equation 6, including MRE, interaction gender continent | -5,573 |
| MAM Equation 6, including MRE, interaction urban continent | -5,579 |
| MAM Equation 6, including MRE, varying coefficient SSA C. Asia S. Asia | -5,972 |
| MAM Equation 6, including MRE, varying coefficient SSA Central Asia South Asia, omitting insignificant varying coefficients | -5,950 |

Source: DHS; calculation by authors.

C. Varying and random effects main model

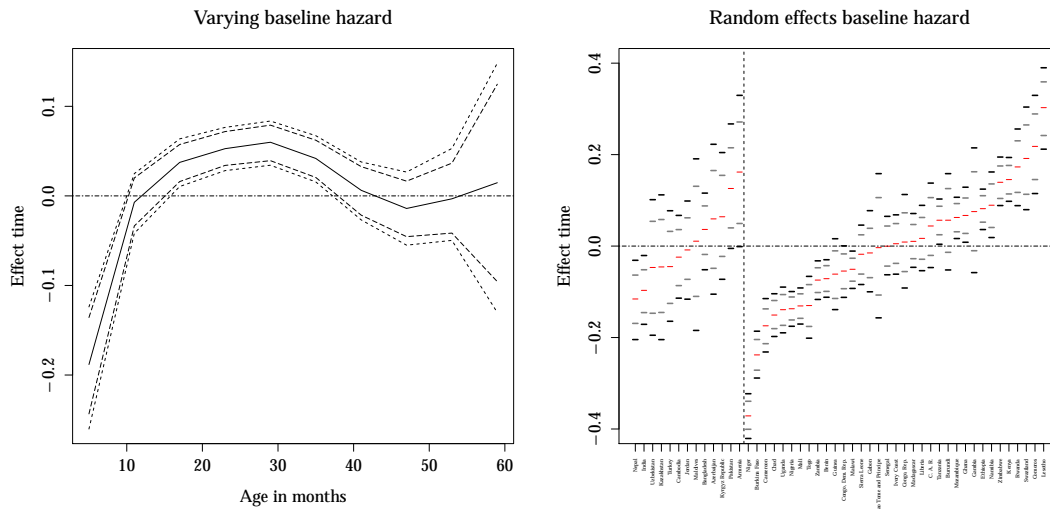


Fig. C.1: Non-linear varying effect of the baseline hazard at level-1. Shown is the with *continent* varying effect, together with 95% and 80% simultaneous credible intervals. Additionally, the multiplicative country random effects are shown. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

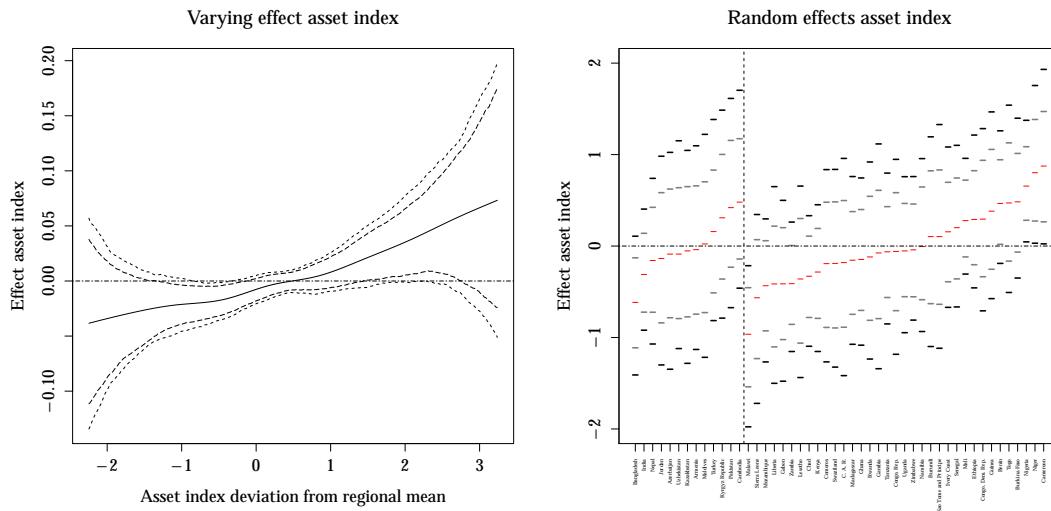


Fig. C.2: Non-linear varying effect of the asset index at level-1. Shown is the with *continent* varying effect, together with 95% and 80% simultaneous credible intervals. Additionally, the multiplicative country random effects are shown. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

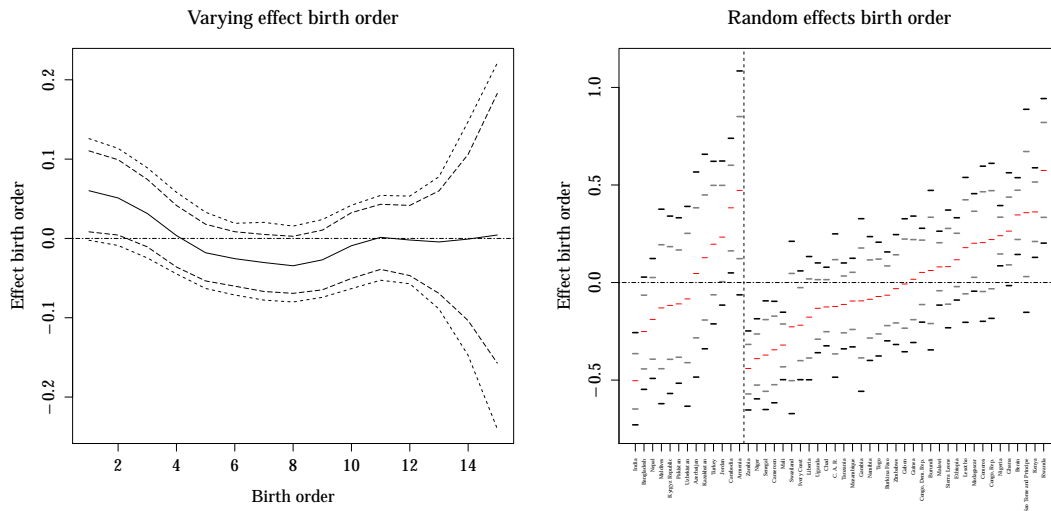


Fig. C.3: Non-linear varying effect of the birth order within the household at level-1. Shown is the with *continent* varying effect, together with 95% and 80% simultaneous credible intervals. Additionally, the multiplicative country random effects are shown. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

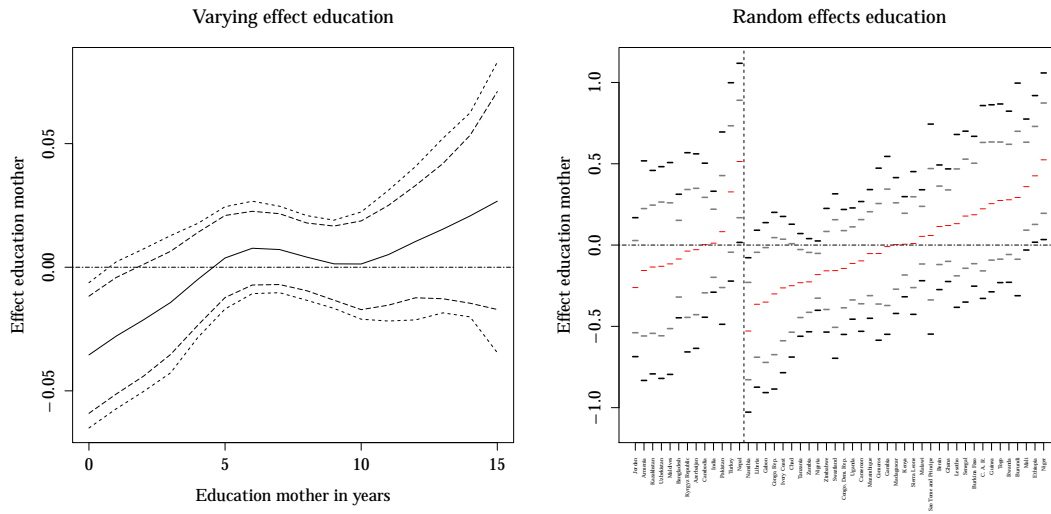


Fig. C.4: Non-linear varying effect of the years of education of the mother at level-1. Shown is the with *continent* varying effect, together with 95% and 80% simultaneous credible intervals. Additionally, the multiplicative country random effects are shown. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

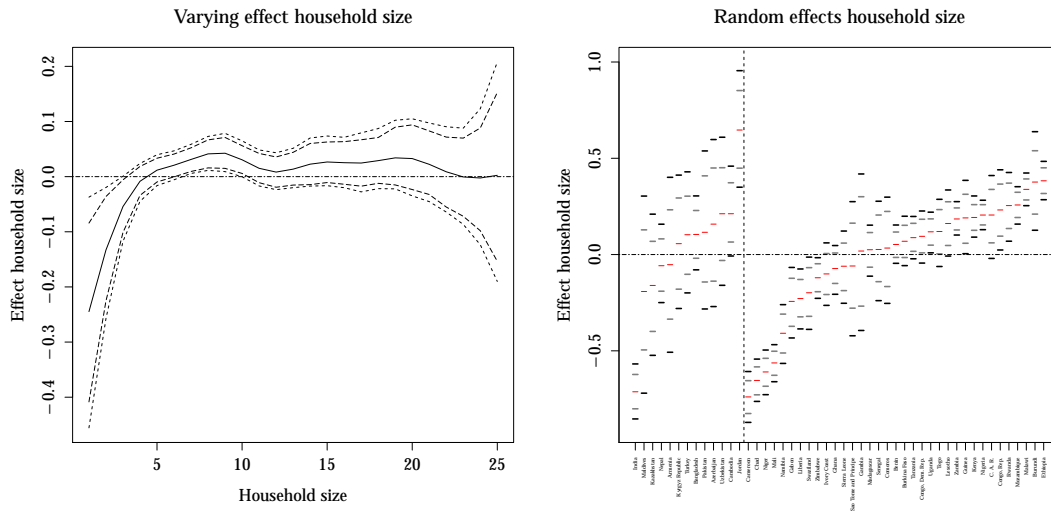


Fig. C.5: Non-linear varying effect of the household size at level-1. Shown is the with *continent* varying effect, together with 95% and 80% simultaneous credible intervals. Additionally, the multiplicative country random effects are shown. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

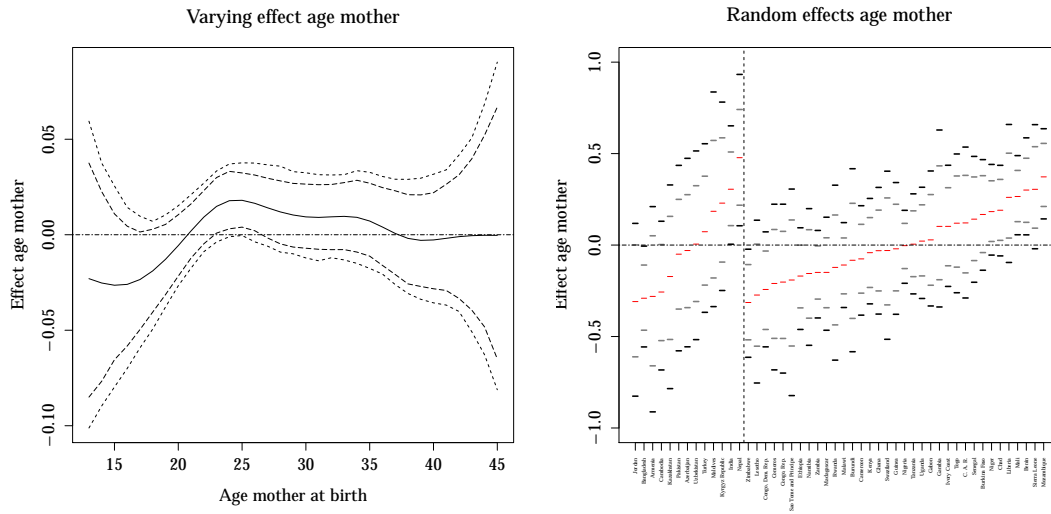


Fig. C.6: Non-linear varying effect of the age of the mother at birth at level-1. Shown is the with *continent* varying effect, together with 95% and 80% simultaneous credible intervals. Additionally, the multiplicative country random effects are shown. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

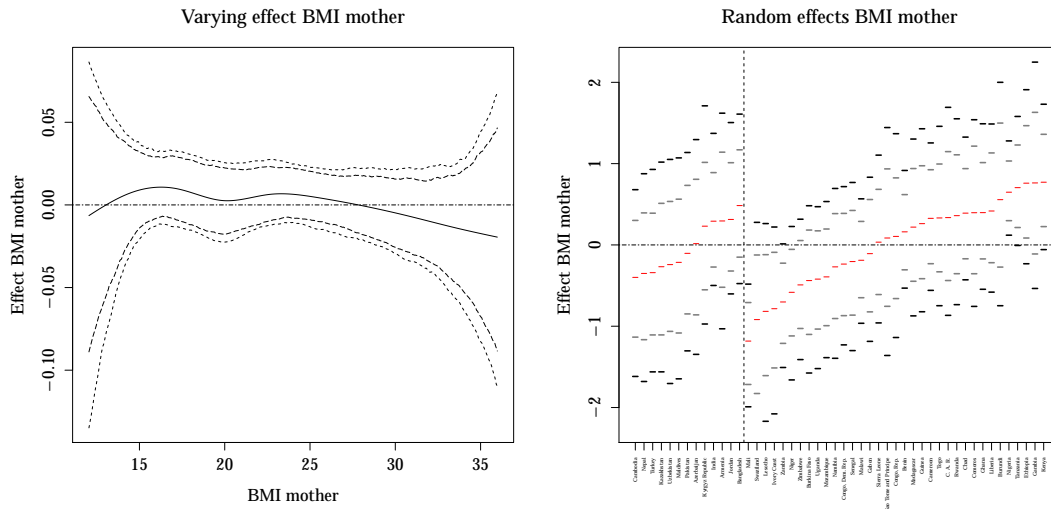


Fig. C.7: Non-linear varying effect of the BMI of the mother at level-1. Shown is the with *continent* varying effect, together with 95% and 80% simultaneous credible intervals. Additionally, the multiplicative country random effects are shown. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

D. Results distinguishing between C./W. and S. Asia

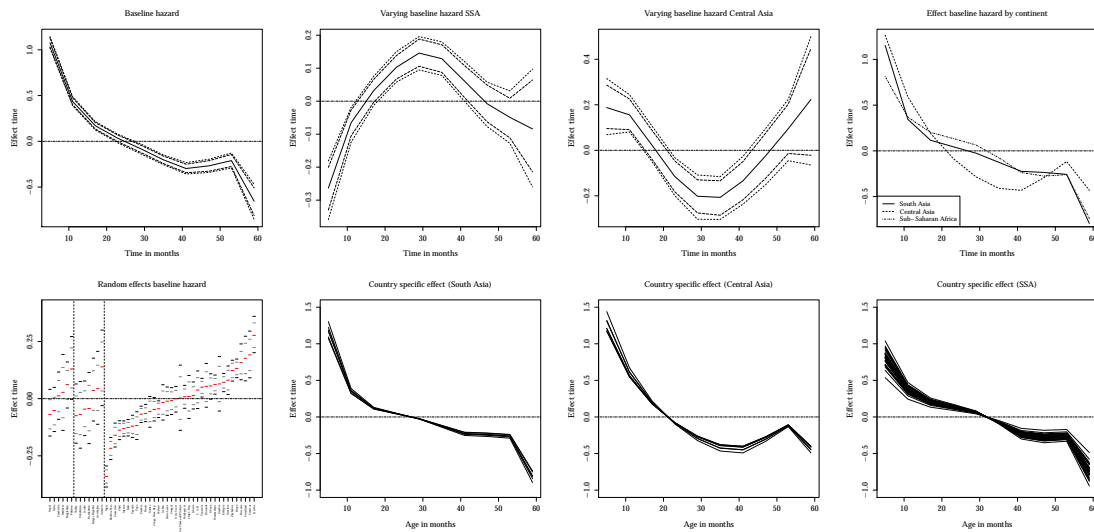


Fig. D.1: Non-linear effect of the baseline hazard at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. South Asian and Western/Central Asian countries are shown to the left of the vertical dashed lines, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

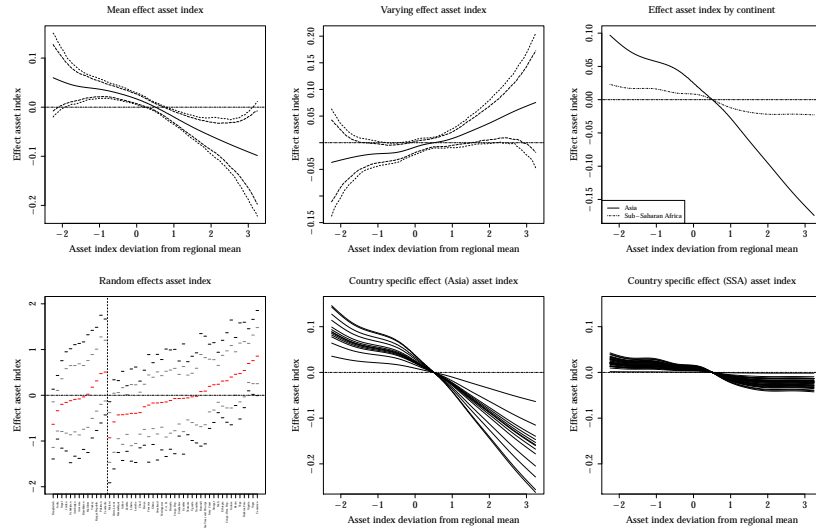


Fig. D.2: Non-linear effect of the asset index at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

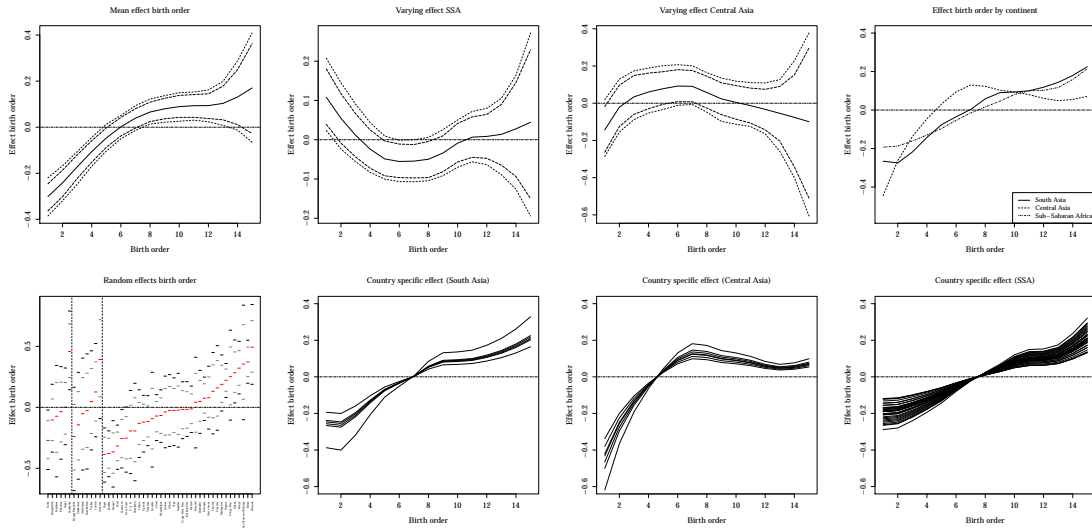


Fig. D.3: Non-linear effect of the birth order within the household at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. South Asian and Western/Central Asian countries are shown to the left of the vertical dashed lines, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

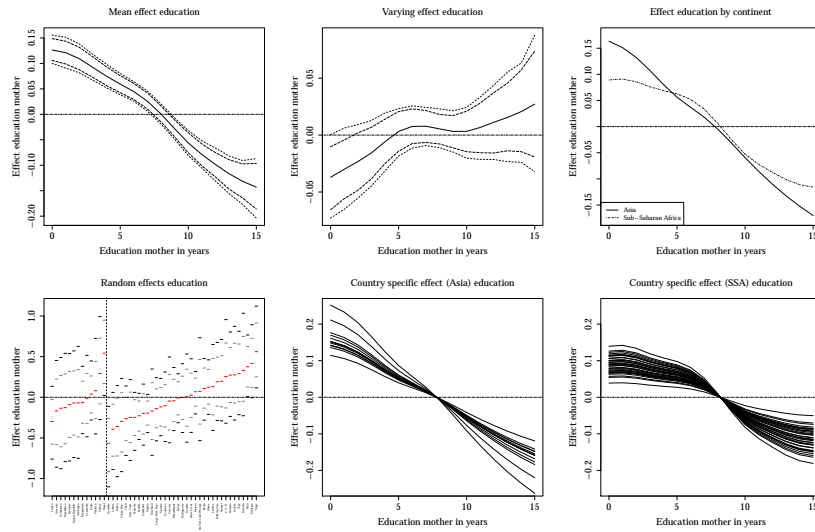


Fig. D.4: Non-linear effect of the mother’s years of education at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

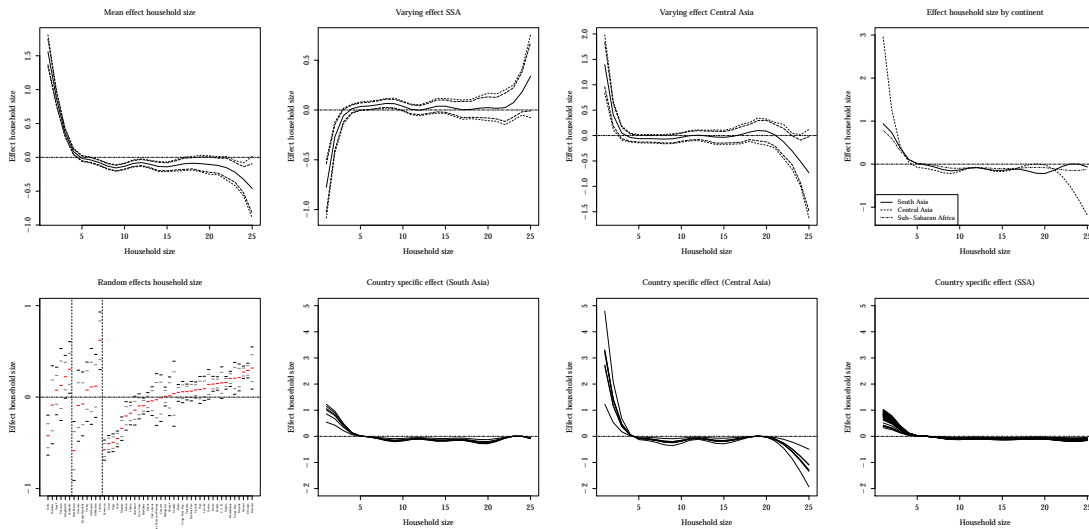


Fig. D.5: Non-linear effect of the size of the household at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. South Asian and Western/Central Asian countries are shown to the left of the vertical dashed lines, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

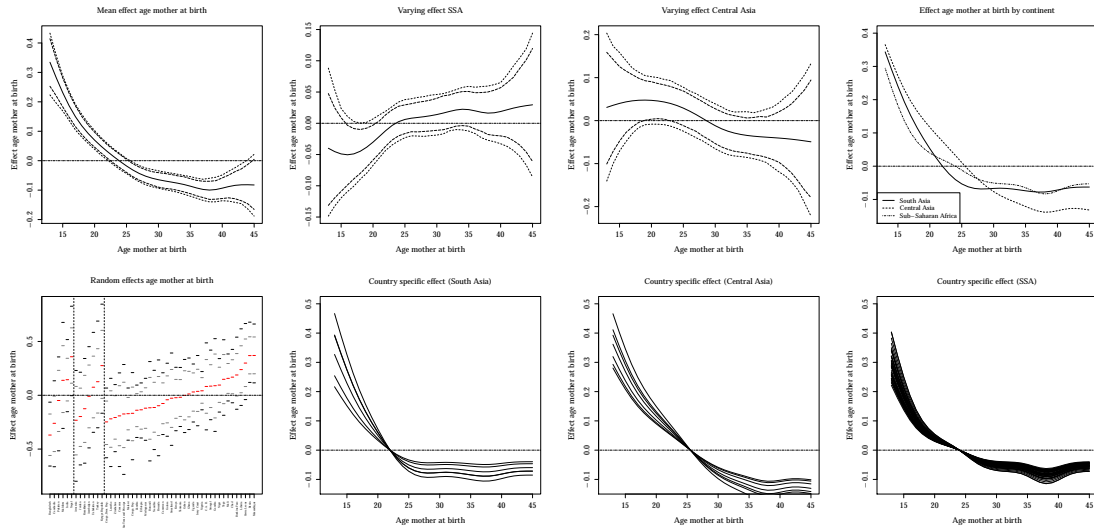


Fig. D.6: Non-linear effect of the age of the mother at birth at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. South Asian and Western/Central Asian countries are shown to the left of the vertical dashed lines, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

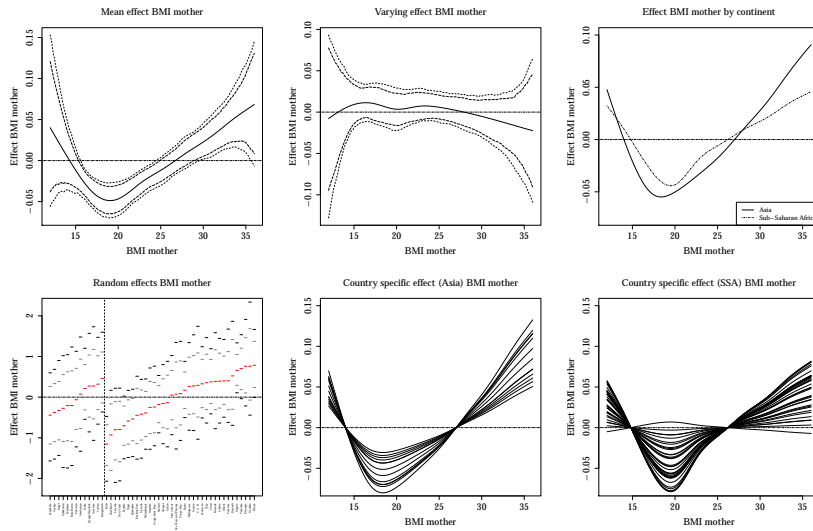


Fig. D.7: Non-linear effect of the mother's BMI at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. Asian countries are shown to the left of the vertical dashed line, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

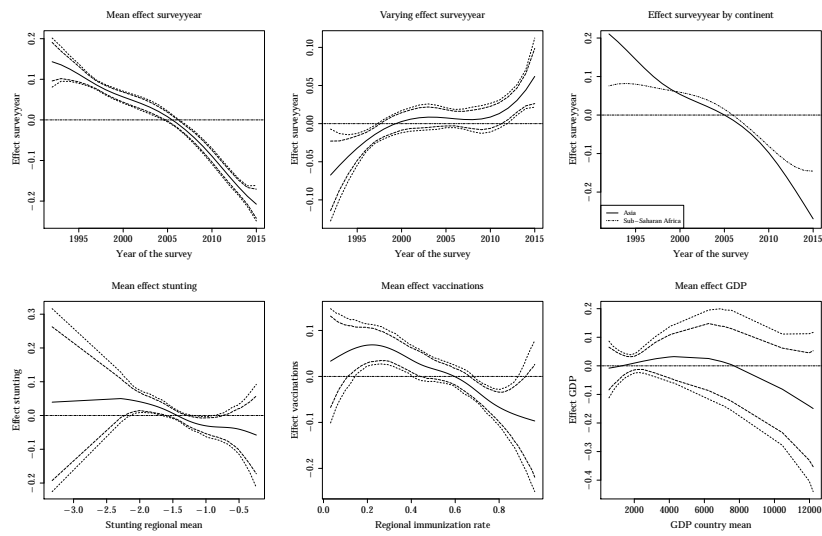


Fig. D.8: Non-linear effects of the survey year at level-1, the nutritional status, and regional immunization rate at level-2, and GDP at level-3. Shown are the mean effects, and for the surveyyear the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect of the geographic regions. **Source:** DHS, and Penn World Table 9.0; calculations by the authors.

E. Results including preceding birth interval

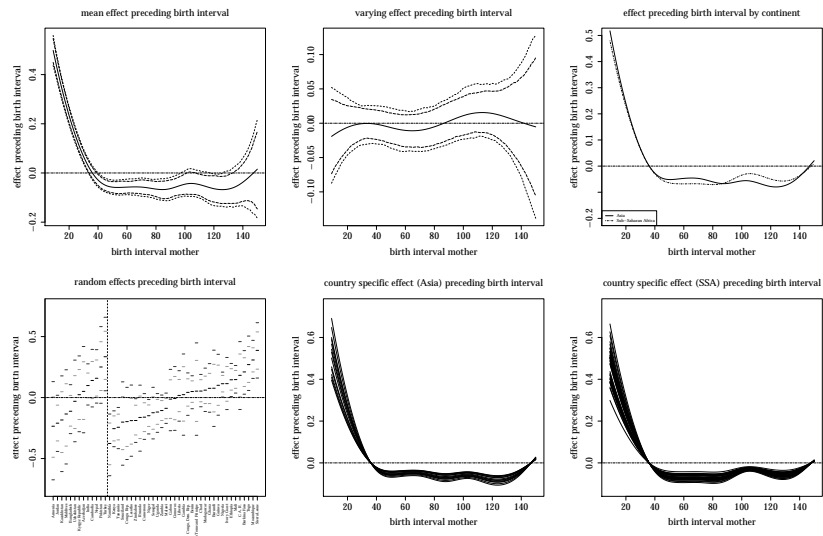


Fig. E.1: Non-linear effect of the mother's preceding birth interval at level-1. Shown is the mean effect, the with *continent* varying effect, both together with 95% and 80% simultaneous credible intervals, and the individual effect for the geographic regions. Additionally, the continent specific mean effect are scaled with the multiplicative country specific random effects. South Asian and Western/Central Asian countries are shown to the left of the vertical dashed lines, while Sub-Saharan African countries are on the right. **Source:** DHS; calculations by the authors.

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

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

- 2017-15 **Kenneth Harttgen, Stefan Lang, Judith Santer, Johannes Seiler:** Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa
- 2017-14 **Christoph Eder, Martin Halla:** Economic origins of cultural norms: The case of animal husbandry and bastardy
- 2017-13 **Thomas Kneib, Nikolaus Umlauf:** A Primer on Bayesian Distributional Regression
- 2017-12 **Susanne Berger, Nathaniel Graham, Achim Zeileis:** Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R
- 2017-11 **Natalia Danzer, Martin Halla, Nicole Schneeweis, Martina Zweimüller:** Parental leave, (in)formal childcare and long-term child outcomes
- 2017-10 **Daniel Muller, Sander Renes:** Fairness views and political preferences - Evidence from a large online experiment
- 2017-09 **Andreas Exenberger:** The Logic of Inequality Extraction: An Application to Gini and Top Incomes Data
- 2017-08 **Sibylle Puntischer, Duc Tran Huy, Janette Walde, Ulrike Tappeiner, Gottfried Tappeiner:** The acceptance of a protected area and the benefits of sustainable tourism: In search of the weak link in their relationship
- 2017-07 **Helena Fornwagner:** Incentives to lose revisited: The NHL and its tournament incentives
- 2017-06 **Loukas Balafoutas, Simon Czermak, Marc Eulerich, Helena Fornwagner:** Incentives for dishonesty: An experimental study with internal auditors
- 2017-05 **Nikolaus Umlauf, Nadja Klein, Achim Zeileis:** BAMLSS: Bayesian additive models for location, scale and shape (and beyond)
- 2017-04 **Martin Halla, Susanne Pech, Martina Zweimüller:** The effect of statutory sick-pay on workers' labor supply and subsequent health
- 2017-03 **Franz Buscha, Daniel Müller, Lionel Page:** Can a common currency foster a shared social identity across different nations? The case of the Euro.

- 2017-02 **Daniel Müller:** The anatomy of distributional preferences with group identity
- 2017-01 **Wolfgang Frimmel, Martin Halla, Jörg Paetzold:** The intergenerational causal effect of tax evasion: Evidence from the commuter tax allowance in Austria
- 2016-33 **Alexander Razen, Stefan Lang, Judith Santer:** Estimation of spatially correlated random scaling factors based on Markov random field priors
- 2016-32 **Meike Köhler, Nikolaus Umlauf, Andreas Beyerlein, Christiane Winkler, Anette-Gabriele Ziegler, Sonja Greven:** Flexible Bayesian additive joint models with an application to type 1 diabetes research
- 2016-31 **Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Simultaneous ensemble post-processing for multiple lead times with standardized anomalies
- 2016-30 **Alexander Razen, Stefan Lang:** Random scaling factors in Bayesian distributional regression models with an application to real estate data
- 2016-29 **Glenn Dutcher, Daniela Glätzle-Rützler, Dmitry Ryvkin:** Don't hate the player, hate the game: Uncovering the foundations of cheating in contests
- 2016-28 **Manuel Gebetsberger, Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Tricks for improving non-homogeneous regression for probabilistic precipitation forecasts: Perfect predictions, heavy tails, and link functions
- 2016-27 **Michael Razen, Matthias Stefan:** Greed: Taking a deadly sin to the lab
- 2016-26 **Florian Wickelmaier, Achim Zeileis:** Using recursive partitioning to account for parameter heterogeneity in multinomial processing tree models
- 2016-25 **Michel Philipp, Carolin Strobl, Jimmy de la Torre, Achim Zeileis:** On the estimation of standard errors in cognitive diagnosis models
- 2016-24 **Florian Lindner, Julia Rose:** No need for more time: Intertemporal allocation decisions under time pressure
- 2016-23 **Christoph Eder, Martin Halla:** The long-lasting shadow of the allied occupation of Austria on its spatial equilibrium
- 2016-22 **Christoph Eder:** Missing men: World War II casualties and structural change
- 2016-21 **Reto Stauffer, Jakob Messner, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Ensemble post-processing of daily precipitation sums over complex terrain using censored high-resolution standardized anomalies *published in Monthly Weather Review*

- 2016-20 **Christina Banner, Eberhard Feess, Natalie Packham, Markus Walzl:** Incentive schemes, private information and the double-edged role of competition for agents
- 2016-19 **Martin Geiger, Richard Hule:** Correlation and coordination risk
- 2016-18 **Yola Engler, Rudolf Kerschbamer, Lionel Page:** Why did he do that? Using counterfactuals to study the effect of intentions in extensive form games
- 2016-17 **Yola Engler, Rudolf Kerschbamer, Lionel Page:** Guilt-averse or reciprocal? Looking at behavioural motivations in the trust game
- 2016-16 **Esther Blanco, Tobias Haller, James M. Walker:** Provision of public goods: Unconditional and conditional donations from outsiders
- 2016-15 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Predictive bookmaker consensus model for the UEFA Euro 2016
- 2016-14 **Martin Halla, Harald Mayr, Gerald J. Pruckner, Pilar García-Gómez:** Cutting fertility? The effect of Cesarean deliveries on subsequent fertility and maternal labor supply
- 2016-13 **Wolfgang Frimmel, Martin Halla, Rudolf Winter-Ebmer:** How does parental divorce affect children's long-term outcomes?
- 2016-12 **Michael Kirchler, Stefan Palan:** Immaterial and monetary gifts in economic transactions. Evidence from the field
- 2016-11 **Michel Philipp, Achim Zeileis, Carolin Strobl:** A toolkit for stability assessment of tree-based learners
- 2016-10 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** Affirmative action or just discrimination? A study on the endogenous emergence of quotas published in *Journal of Economic Behavior and Organization*
- 2016-09 **Loukas Balafoutas, Helena Fornwagner:** The limits of guilt
- 2016-08 **Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Spatial ensemble post-processing with standardized anomalies
- 2016-07 **Reto Stauffer, Jakob W. Messner, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Spatio-temporal precipitation climatology over complex terrain using a censored additive regression model
- 2016-06 **Michael Razen, Jürgen Huber, Michael Kirchler:** Cash inflow and trading horizon in asset markets
- 2016-05 **Ting Wang, Carolin Strobl, Achim Zeileis, Edgar C. Merkle:** Score-based tests of differential item functioning in the two-parameter model

- 2016-04 **Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Non-homogeneous boosting for predictor selection in ensemble post-processing
- 2016-03 **Dietmar Fehr, Matthias Sutter:** Gossip and the efficiency of interactions
- 2016-02 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Rankings and risk-taking in the finance industry
- 2016-01 **Sibylle Puntcher, Janette Walde, Gottfried Tappeiner:** Do methodical traps lead to wrong development strategies for welfare? A multilevel approach considering heterogeneity across industrialized and developing countries

University of Innsbruck

Working Papers in Economics and Statistics

2017-15

Kenneth Harttgen, Stefan Lang, Judith Santer, Johannes Seiler

Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa

Abstract

Despite improvements in global child health within the last three decades, under-5 mortality remains significantly high in Sub-Saharan and Asia. Both regions did not achieve the MDG target of reducing under-5 mortality by two thirds by 2015. The underlying causes of under-5 mortality differ significantly between countries and between regions, which highlights the need to expand our understanding of the determinants of child health in developing countries. By comparing the two geographic regions of the world with the highest under-5 mortality rates, we aim to gain new insights, and bring out potential differences between these regions and the causes of under-5 mortality. In addition, we aim to identify non-linear relationships between under-5 mortality and specific explanatory variables. We analyze a large data set consisting of 35 Sub-Saharan-African countries, and 13 Asian countries, using a multilevel discrete time survival model that takes advantage of a recently developed multilevel framework with structured additive predictor in a Bayesian setting. We analyze data from 131 individual surveys from 1992 to 2015, allowing for potential non-linear effects and cluster specific heterogeneity within models. We find strong non-linear effects for the baseline hazard, the household size, the age of the mother, the BMI of the mother, and the birth order of the child. Additionally, we find considerable differences in determinants between Asian and Sub-Saharan Asian countries.

ISSN 1993-4378 (Print)

ISSN 1993-6885 (Online)