Selective mortality and undernutrition in low- and middle-income countries

Kenneth Harttgen, Stefan Lang, Johannes Seiler

Working Papers in Economics and Statistics

2017-27

University of Innsbruck
http://uibk.ac.at/eeecon/

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 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

Selective mortality and undernutrition in low- and middle-income countries

Johannes Seiler[‡], Kenneth Harttgen[†], and Stefan Lang[‡]

[†]ETH Zurich, Department of Humanities, Social and Political Sciences, Clausiusstr. 37, 8092 Zurich, Switzerland; E-mail: harttgen@nadel.ethz.ch; Phone: +41 44 632 9825
[‡]University of Innsbruck, Department of Statistics, Universitätsstr. 15, 6020 Innsbruck, Austria

27th July 2018

ABSTRACT Anthropometric indicators, in particular, the height-for-age z-score, are found to be lowest in South Asia, compared to other geopolitical regions. Despite a close relationship between undernutrition and mortality rates, the highest mortality rates are concentrated in sub-Saharan Africa, whereas the lowest anthropometric indicators are concentrated in South Asia. Accounting for this survival bias, i.e. selective mortality, this discrepancy between the anthropometric indicators between South Asia and sub-Saharan Africa should be expected to decrease. Individual data of children from six waves of Demographic and Health Surveys is used to assess the impact of selective mortality on the standard anthropometric indicators, such as stunting, underweight and wasting, for a large cross-section of 35 low- and middle-income countries between 1991 and 2015. Using a matching approach, we estimate values for the anthropometric status of deceased children. As robustness check the anthropometric status of deceased children is predicted using semiparametric regression. The results are twofold: first, both methods reveal that the imputed values for the anthropometric indicators are significantly lower compared to the observed anthropometric indicators. Second, since the share of deceased children in our sample is below 10%, accordingly, the contribution of the anthropometric status of deceased children to the overall anthropometric indicators are small and the overall anthropometric indicators are only altered marginally.

KEYWORDS child mortality, malnutrition, selective mortality, low- and middle-income countries, comparative country studies

JEL CLASSIFICATION: I15, I32, J13, O57

1. INTRODUCTION

Since the formulation of the Millennium Development Goals (MDG) low- and middle-income countries have made significant progress in reducing mortality rates, as well as the reduction in the prevalence of undernutrition among children (You et al., 2015a,b; Black et al., 2013). Highest mortality rates are observed in sub-Saharan Africa, whereas, in particular, the prevalence of stunting if found to be to be higher for children in Asian countries (Klasen, 2008; Paciorek et al., 2013; You et al., 2015b; Jayachandran & Pande, 2017), some papers refer to this observation as so-called South Asian enigma (for instance, Ramalingaswami et al., 1996). Besides this, several authors (for example, Pelletier et al., 1995; Pelletier & Frongillo, 2003; Bryce et al., 2005; Black et al., 2008, 2013) find a strong and positive relationship between the nutritional intake of children and their mortality risk, and around 50% of all deceased children younger than five can be attributed to undernutrition.

The results of Pelletier et al. (1995) are based on the analysis of 53 developing countries and find that around 56% of the observed deaths can be attributed to undernutrition. Bryce et al. (2005) concludes similarly and finds that undernutrition is responsible for around 53% of all deaths of children up to the age of five years. Both studies, however, use country aggregates to estimate the effect of undernutrition on the risk of dying, and present estimates for countries, respectively, geopolitical regions, based on aggregated data. This is arguably less precise in comparison to individual data.

For the analysis of the relationship of malnutrition and mortality on a small scale, ideally, clinical data should be used. Data from hospitals is scarce and as far as we are aware only one such study exists by Sachdeva et al. (2016). The analysis in this study is based on hospitalized children between six and 59 months. The observational data for the three standard z-scores are lower for deceased children compared to living children.

Assuming the relationship between undernutrition and mortality to be correct, after accounting for selective mortality, we would expect that undernutrition rates in sub-Saharan Africa are much higher and the discrepancy between the undernutrition rates between South Asia and sub-Saharan Africa would decrease. Due to the close relationship of undernutrition and mortality, potentially a sampling bias is introduced which is also called selective mortality. Causing the observed anthropometric status not correctly reflecting the unobservable anthropometric status of deceased children. Differences in the development of children in the first years of life can be attributed, to a large extent, to environmental factors, and the impact of genetics and ethnicity is negligible (Martorell & Habicht, 1986; Jorde & Wooding, 2004; WHO Multicentre Growth Reference Study Group & de Onis, 2006; Grantham-McGregor et al., 2007). Therefore, the observed differences between Asia and sub-Saharan Africa can only be attributed to a minor extent to the latter two factors.

The effect of selective mortality was already analysed by several authors, for instance, Boerma et al. (1992); Pitt (1997); Dancer et al. (2008); Alderman et al. (2011). The first and second articles examine 17 and 13 surveys, respectively, from low- and middle income countries. Both articles conclude that even though mortality is positively correlated with the nutritional status, the overall anthropometric status is only influenced marginally when accounting for a selection bias. Dancer et al. (2008) emphasize for data from Bangladesh, the close relationship between nutrition and mortality and find that sample selection is not found for boys indicating that the household's resources are not evenly distributed between boys and girls. The most recent study by Alderman et al. (2011) tries to determine the impact of selective mortality on children's developmental status for a sample of Indian children to determine differences of the observed anthropometric status of non-deceased children and the unobserved overall anthropometric status caused by selective mortality. Analysing the link between undernutrition and mortality is potentially biased due to this selection. They used a matching approach to get plausible estimates for the anthropometric status of deceased children that are used to calculate the overall anthropometric status. Besides this line of research, which focuses on determining the effect of selective mortality on undernutrition, in the literature, another line of research exists that analyses the long-term consequences of selective mortality, trying to link childhood undernutrition and adult height (for instance, Moradi, 2010; Gørgens et al., 2012).

By pooling 140 surveys from 35 countries between 1991 and 2015, we assess the effect of selective mortality on the unobserved anthropometric status of deceased children, as well as the overall anthropometric status for the complete sample, individually for the geopolitical regions, and individually on the country level. This is of great importance for the literature, which uses hierarchical models in the analysis of child mortality, (for instance, Boco, 2010; Harttgen et al., 2015; Adedini et al., 2015; Harttgen et al., 2017) and similar future research, as aggregating the nutritional status potentially introduces a bias. We extend the current knowledge of the effect of selective mortality on the anthropometric status, and therefore aim to contribute to the literature in the following regards.

First, we assess whether selective mortality exists and estimate plausible values for the anthropometric indicators of deceased children and the overall anthropometric status.

Second, we examine the impact of selective mortality on undernutrition rates for a sample of low- and middle-income countries from Asia, Latin America, and sub-Saharan Africa. We also determine whether differences between the geopolitical regions exist. We aim to identify whether the effect of selective mortality differs between countries by analysing selective mortality individually for the included countries.

Third, we analyse whether the impact of selective mortality on undernutrition rates changed over time by splitting the sample at the year 2000. To make this feasible, and to make the composition of the effect clear, we restrict our sample to countries which were sampled before and after 2000. The data consists of 140 individual surveys, between 1991 and 2015 from 35 Asian, Latin American, and sub-Saharan African countries.

Fourth, we evaluate if a systematic underestimation of the effect of undernutrition is introduced by the joint analysis of child mortality and undernutrition. Undernutrition and mortality are closely related, which makes it necessary to include the nutritional status in the analysis of mortality. This however, is not possible at the individual level as the anthropometric status are observed by the Demographic and Health Surveys (DHS), only for non-deceased children. A hierarchical model is of particular use in this situation and several studies exist that analyse underlying causes of child mortality using this technique (for instance, Misselhorn & Harttgen, 2006; Adedini et al., 2015; Harttgen et al., 2015, 2017). Aggregating the anthropometric indicators at a regional level can, however, due to the close relationship bias the results.

The results can be summarized as follows, the imputed values for the anthropometric status of deceased children are significantly lower compared to the observed anthropometric indicators. However, since the share of deceased in our sample children is below 10%, accordingly, the contribution of the anthropometric status of deceased children to the overall anthropometric indicator is small, and the overall anthropometric indicators are only altered marginally.

The reminder of this paper is organized as follows: Section 2 will give an overview of the data set we used and will also give descriptive statistics for the anthropometric indicators. Section 3 illustrates the potential impact of selective mortality on the anthropometric indicators. Section 4 will briefly illustrate the applied statistical method. Section 5 and Section 6 present and discuss the results for the anthropometric indicators stemming from this approach, and give concluding remarks to our findings for the anthropometric indicators.

2. DATA AND SUMMARY STATISTICS

The data used in the analysis are sourced from each countries' national DHS. The surveys are conducted by Macro International Inc., Calverton, Maryland in cooperation with local administration and receive funding from United States Agency for International Development (USAID). The 7th wave is currently being conducted, and there is usually a period of five years between waves. The data is standardized across countries and include among others, information on topics such as child health, child mortality, child and maternal education, maternal health, and quality of the household's sanitation and dwelling. Our analysis includes 769,044 children from 35 low- and middle-income countries between 1991 and 2015. To allow for an interpretation over time, we include countries from the three mentioned regions, for which at least one survey before and after 2000 is available. According to the classification made by the DHS seven countries are located in Latin America & Caribbean (henceforth Latin America), four countries are located in South & South-east Asia (henceforth Asia) and 24 countries are located in sub-Saharan Africa.

When comparing the distribution of deaths in the first three years of life, it becomes noticeable that deaths are not evenly distributed over time, instead they are skewed-right. This is also illustrated in Figure 1. Table I shows the share of deceased children grouped by gender and age for the surveys before and after 2000 for the global sample, as well as the regional sub samples. It can be seen that differences between the geopolitical regions exist for both male and female observations. The share of deceased children who die in the first month is lowest for both genders in sub-Saharan Africa (40.5% of boys, 33.7% of girls). This shows that the risk of dying varies with respect to the geographic region, and is not evenly distributed over the age range from 0 to 35 months. The pattern that emerges from this figure is the following: exposure to a high risk of dying for a longer period is predominantly observed for children from sub-Saharan Africa. In other geopolitical regions the share of children dying in the first month of life is between 45% and up to 60%, Asia (60.4% of boys, 53.5% of girls), and Latin America (48.5% of boys, 46.7% of girls) indicating that the exposure to the mortality risk strongly declines after the first months. This pattern can also be seen in Figure 1 where the cumulative hazards by gender and geopolitical region are plotted against time and the curves are flattening with increasing age. This illustrates that in the statistical analysis it is important to account for the asymmetric distribution of deaths over time. With respect to the distribution over time a similar pattern was observed by Alderman et al. (2011) for India.

[Please insert FIGURE 1 about here]

[Please insert TABLE I about here]

Globally, the share of stunted children is decreasing over time; the average z-score for height-for-age increased from -1.66 before 2000, to -1.37 after 2000 for boys, and from -1.47 to -1.18 for girls. See Table II and Table III for a more elaborate analysis. Nevertheless, undernutrition and its resulting effects (diarrhoea, pneumonia, weaker immune system) are still a prevailing cause of child mortality (Grantham-McGregor et al., 2007; Black et al., 2008; Liu et al., 2012, 2015). Large dispersions in the height-for-age z-scores between the three geopolitical regions are observed. The anthropometric indicator for height-for-age is found to be lowest in Asia for both female and male observations, followed by sub-Saharan Africa and Latin America. Using the definition of the WHO for not stunted, moderate stunted and severely stunted, the same pattern emerges across the geopolitical regions. At a global scale the share of non-stunted children increased slightly from 56.3% to 63.5% for boys and 61.6% to 68.8% for girls. While the share of severely stunted children decreased over time. The share of severely stunted children is found to be highest in Asia, however it decreased from 28.3% to 18.9% for boys and from 26.1% to 16.2% for girls. The lowest shares of severely stunted children are found in Latin America. Within geopolitical regions, differences between countries exist, which is illustrated in Table A.1. To account for this pattern in the statistical analysis, in addition to estimating the effect of selective mortality at the global level we will also estimate the effect of selective mortality for the three geopolitical regions separately, as well as for the countries within these regions where stunting, mortality rates, are found to be highest. Previous to the statistical analysis we will consider some simple imputations with thresholds used by Alderman et al. (2011) and found in the literature (for instance, Pelletier et al., 1995).

[Please insert TABLE II about here]

[Please insert TABLE III about here]

3. ILLUSTRATION OF SELECTIVE MORTALITY AND ITS POTENTIAL IMPACT IN ANTHROPOMETRIC INDICATORS

Selective mortality. Since anthropometric indicators of children in most surveys are only observed for living children, a potential bias is introduced in any further analysis by this selection effect, also referred to as selective mortality. Substituting the average overall anthropometric indicator of deceased and non-deceased children (Z_{all}), for instance by aggregating the observed anthropometric indicator (Z_s) at a regional level could potentially yield biased estimates. Since it can be assumed that due to the relationship between undernutrition and mortality, the unobserved average anthropometric status of deceased children (Z_d) is lower as Z_s. When the share of deceased children (P_d) is known the relationship between Z_{all}, Z_s and Z_d can be expressed as follows (Alderman et al., 2011):

$$Z_{all} = Z_s(1 - P_d) + Z_d P_d, \tag{1}$$

A certain share of fatalities can be attributed to undernutrition and therefore it can be assumed that the unobserved anthropometric status of deceased children is lower than the observed anthropometric status which gives the property of $Z_d < Z_s$. Neglecting this property could potentially lead to an overestimation of Z_{all}. For the corresponding period, the under-3 mortality rate decreased from 111 to 83 per 1,000 live births, accordingly the contribution of the observed z-score (Z_s) to Z_{all} is much higher. Thus, Z_{all}, the overall z-score, will only be altered in the case of high mortality rates (P_d) or in the case where the imputed z-score for deceased children (Z_d) is close to the cut-off for outliers, as defined by WHO. To account for this, also comparing Z_d with Z_s will offer interesting insights.

Due to the large share of deaths that can be accounted to undernutrition, the unobserved z-score of deceased children should be lower compared to the observed z-score of living children. Making a statement about the size of the effect is rather difficult. Simple imputations carried out in the next section should give new insights about the size of the effect.

Simple imputations demonstrating the potential influence of selective mortality. To illustrate the potential influence of selective mortality on the overall z-score (Z_{all}), and to compare our results with the estimates of Alderman et al. (2011), we impute values for the unobserved anthropometric indicators of deceased children under the counter factual outcome. The imputation is done analogous to Alderman et al. (2011) and the same four input scenarios are considered: first (I1), imputing for the unobserved anthropometric status of deceased children, values which result in a statistical significant difference of the observed and overall anthropometric status ($\alpha = 0.05$). Second (I2), imputing for the anthropometric status of deceased children as double that of the observed average anthropometric status of living children. Third (I3), the lower bounds considered by the WHO as cut-offs to identify outliers are imputed[1]. And fourth (I4), the upper bound for the share of deceased children, which can be attributed to undernutrition found by Pelletier et al. (1995) are imputed. Based on this, around two thirds of the share of deceased children can be traced back to undernutrition.

All four imputation scenarios are conducted for the height-for-age, weight-for-age, and weight-for-height z-scores, however only the results for the height-for-age z-scores are shown. These imputations have in common that the unobserved overall z-score would decrease in all four scenarios. The differences between boys and girls are only moderate, with girls generally having a higher anthropometric indicator. Depending on the scenario the overall z-score would decrease by up to approximately 30%. Imputing the (new) lower-bounds recommended by the WHO as the cut-off for outliers yields the largest change of the overall z-score. These are somehow similar values compared to the estimated values by Alderman et al. (2011), even though we used the new lower-bounds as cut-off in one imputation scenario. This extreme case yields the highest changes for the overall z-score. Table IV shows the detailed results under the four imputation scenarios.

[Please insert Table IV about here]

4. MODELLING THE ANTHROPOMETRIC STATUS OF DECEASED CHILDREN

The previous section shows the changes of the anthropometric status without considering individual characteristics and illustrates that the overall anthropometric status will only be altered

---

[1] Contrary to Alderman et al. (2011) who are using the old cut-off for identifying outliers defined in WHO Expert Committee (1995), we use the lower bounds recommended by WHO (2010), since DHS data use this cut-off points for identifying outliers. The cut-off points (old cut-off in brackets) for the anthropometric indicators are:
Height-for-age (stunting): < -6.0 (-5.0) and $> +6.0$ ($+3.0$);
Weight-for-age (underweight): < -6.0 (-5.0) and $> +5.0$ ($+5.0$);
Weight-for-height (wasting): < -5.0 (-4.0) and $> +5.0$ ($+5.0$).

in some specific cases. This emphasizes a better understanding of the anthropometric status of deceased children.

4.1. Matching Approach

The survival status depends to a large extent on individual characteristics, characteristics of the mother, general living conditions (Mosley & Chen, 1984; Schultz, 1984) and unobservable factors. The survival model needs to be able to account for frailty in the children's health outcome. Assuming a Weibull proportional hazard model as in Alderman et al. (2011), the hazard function at time t is given by,

$$h_(t) \quad = \gamma \lambda t^{\gamma - 1}, \tag{2}$$

with the parametrization in the proportional hazard form of $\lambda = \exp(\mathbf{x_i}\beta)$. Whereas the survival function S_t is given by,

$$S(t) = \exp(-\lambda t)^{\gamma}. \tag{3}$$

Including in the Weibull proportional hazard model a term to control for unobserved heterogeneity, a so-called frailty term α, allows us to control for unobserved factors in the children's health outcome. Assuming the unobserved factor to be inverse-Gaussian distributed, the hazard function can be written as follows (Gutierrez, 2002; Munda et al., 2012),

$$h_i(t|\alpha) \quad = \alpha h_0(t) \exp(\mathbf{x_i}\beta). \tag{4}$$

The regression parameters of the Weibull proportional hazard model can be estimated using maximum likelihood procedures. The matching score will be derived from these estimates, which allows values to be assigned to the unobserved anthropometric indicators of deceased children by using the observed anthropometric indicators of non-deceased children. This allows potential differences to be determined in the anthropometric indicators of deceased children, and the observable anthropometric indicators of living children. Matching follows the ideas of Rubin (1976, 1979); Rosenbaum & Rubin (1983), who introduced propensity score matching. The matching score is not derived from a probit or logit model, instead it is derived from the predictions made by the Weibull proportional hazard regression. The advantages are twofold: first, it is capable to model asymmetric distributions, which is important since deaths are not evenly distributed over time, which is illustrated in Figure 1 and Table I. Second, survival models allow for censored and truncated observations to be included. The matching score is then used to find pairs of observations from the two groups that have a similar matching score via a nearest neighbour algorithm. After the construction of suitable pairs, the observed z-score of the vital child i is assign to the deceased child j. The algorithm only assumes, as pointed out by Heckman et al. (1999) that observations in both samples have common support, ensuring sufficiently large overlap between both groups. Before showing the results of this imputation, descriptive statistics of variables that will be used are shown. We included observations up to the third year of living to make our estimates at a country and regional level comparable to the results for the Indian sample of Alderman et al. (2011).

4.2. Modelling the anthropometric status

Using a regression framework and predicting the anthropometric status of deceased children can be seen as alternative approach to assess the difference between observed and unobserved z-scores. To gain a maximum of information from the data, semiparametric regression is applied, and the usually assumed linear predictor η_i^{lin} of the response is substituted by the more general additive predictor η_i^{add}. Using semiparametric regression allows us to approximate the metric covariates by penalized splines. The corresponding semiparametric model can be written as follows (for a more thorough description of the estimation see, for instance Fahrmeir et al., 2013, Chapter 7 and Chapter 8):

$$\begin{aligned} \text{z-score}_i &= f(z_{i1}) + ... + f(z_{iq}) + \gamma_0 + \gamma_1 x_{i1} + ... + \gamma_k x_{ik} + \epsilon_i \\ &= f(z_{i1}) + ... + f(z_{iq}) + \eta_i^{lin} + \epsilon_i \\ &= \eta_i^{add} + \epsilon_i. \end{aligned} \tag{5}$$

As the nutritional status of the child (z-score) is measured continuously we assume the response to be normally distributed with variance σ^2 ($\mathcal{N}(\mu_i; \sigma^2)$). The response is modelled as a function of socio-economic, maternal and child specific characteristics of undernutrition following closely the conceptual framework of malnutrition developed by the UNICEF (UNICEF, 1990). In recent years this framework has been applied to model socio-economic, maternal and child specific characteristics of undernutrition (for instance Fenske et al., 2013; Lang et al., 2014). $f(z_{i1})$ to $f(z_{iq})$ represent the smooth functions of the included metric covariates: the household's asset index, the number of household members, the age of the mother, vaccination coverage, the year the survey was conducted and in order to account for time trends. In addition, we include country random effects. Then x_{i1} to x_{iq} represent the included categorical covariates that enter the model linearly: the gender of the child, the educational level of the mother and an indicator whether the household is located in urban or rural area.

The obtained estimation results will then be used to predict the z-scores of deceased children using out-of-sample prediction and to impute the predicted z-scores to the unobserved z-scores of deceased children. The estimation is carried out using the **R**-package **mgcv** (Wood, 2011).

In addition, to allow the effects of the metric covariates to vary between geopolitical regions, we include varying coefficients for the metric covariates. This type of expansion to semiparametric models was first introduce by Hastie & Tibshirani (1993). The idea is to allow an interaction between the metric covariate z_i and a categorical covariate x_i. The extension of η_i^{add} by the varying coefficient can be specified in its simplest form (the categorical covariate is binary) as follows:

$$\eta_i^{add} = f_1(z_{i1}) + ... + f_q(z_{iq}) + f_{(z_1|x_1)}(z_{i1})(x_{i1}) + ... + \boldsymbol{x}_i'\boldsymbol{\gamma}. \tag{6}$$

This allows the effect of the categorical covariate to vary smoothly over the range of the metric covariate.

5. RESULTS

5.1. Matching approach

Descriptive statistics of the explanatory variables are shown in Table V, for both the complete sample, the regional sub-samples and the periods before and after 2000. Throughout all regions the share of female observations is always slightly below 50%. The average household size decreased in all regions when comparing before and after 2000, as well as the share of mothers without any education. The share of mothers with no formal education decreased and the average years of education of mothers increased slightly. The average size of the households remains almost constant over time. Most of the observations are located in sub-Saharan Africa, which also reflects that most developing countries are located in sub-Saharan Africa. Additionally, the descriptive statistic reveals that the average z-scores increased over time, indicating a better nutritional situation in general (see Table II and Table III). We find the results to be similar between the three different anthropometric indicators and only the results for stunting will be presented.

[Please insert TABLE V about here]

Results parametric hazard regression Table VI and Table VII show the results of the Weibull proportional hazard model. The predictions obtained from the Weibull proportional hazard model are then used as propensity score to assign the anthropometric indicator of non-deceased children to the unobserved anthropometric indicator of deceased children. The Weibull proportional hazard model includes characteristics of the children, the mother and the household. The DHS do not report a subjective birth weight measure[2] of the mother and information for the child's birth weight is not available for all observations, even though it is an important indicator, especially for the health status of new-borns (Claeson et al., 2000; Alderman et al., 2011). Including the child's birth weight would have reduced the sample size by around one third. More importantly this would have increased on average the height-for-age anthropometric indicator by around 0.4 standard deviations, which will introduce a bias in the estimates. This effect will be captured indirectly through including a frailty term.

Our estimates confirm the estimated relationships between child mortality and the included covariates already pointed out, among others, by Alderman et al. (2011). Most of our estimates are similar in magnitude and shape. Controlling for gender, our results show that girls have higher survival chances than boys. While this initially seems to be in contrast to the results of Alderman et al. (2011), where they find an insignificant coefficient for gender, we accounted for this discrepancy. This is because their study is done for India, where the existing preference for male children counter balances this effect and mortality rates are lower for girls than for boys in countries where gender preferences are insignificant (Arnold, 1997). In addition to the gender, we find that children living in wealthier households have a higher survival probability as well children whose mothers have secondary or higher education. This is in agreement with the findings of Cleland & Van Ginneken (1988); Vollmer et al. (2017). The reason for this is twofold: first, it can be assumed that mothers with a higher educational level are more likely to find better paying jobs. Second, they are more capable of assessing relevant information, which can be potentially beneficial to the children's health status. Children in urban areas, probably due to better access to sanitation, and/or drinking water, have also better survival outcomes compared to children born in rural areas, which is a quite common result found in the literature (Harttgen & Günther, 2011; Harttgen et al., 2015, 2017).

[Please insert TABLE VI about here]

[Please insert TABLE VII about here]

---

[2] The birth weight of the children was assessed by the mother as either below average, average or above average.

Pooling low- and middle-income countries by geopolitical region. Figure 2 and Figure 3 illustrate the changes for the overall height-for-age z-score in comparison to the observed z-score of the complete sample, as well as of the three geopolitical regions, before and after 2000 for boys and girls, respectively. The left-hand panel depicts the observed z-scores, the constructed overall z-scores and imputed z-scores of deceased children. Whereas the right-hand panel shows the differences between the observed z-scores, the constructed overall z-scores and the differences between the observed z-scores and the imputed z-scores. The observed z-scores and the constructed overall z-scores are more or less identical and the difference is never statistically significant. This lies within our expectations, as the constructed overall z-scores are the aggregate of the observed z-scores and the imputed z-scores, and with an observed under-3 mortality rate of around 100 per 1,000 live birth, the observed z-scores have a much higher weight in comparison to the imputed z-scores. Significant changes could only be expected with an increased share of deceased children, or a large amount of the imputed z-scores being close to the cut-off for outliers as defined by the WHO. In addition, differences across geopolitical regions and gender for the observed z-scores and constructed overall z-scores are negligible.

It is more interesting and important to compare the observed z-scores of non-deceased children with the imputed z-scores of deceased children, which is depicted in the right-hand panels of Figure 2 and Figure 3. Comparing the observed z-scores with the imputed z-scores reveals for the global sample that the non-observable z-scores of deceased children should be lower in comparison to the observed z-score of non-deceased children. In the period before 2000 the difference is around 0.05 for both boys and girls. To our surprise, the difference increases to around 0.075 for boys and around 0.1 for girls after 2000. We also find that the difference varies across geopolitical regions. For South Asia the differences before and after 2000 between the observed z-score and the imputed z-score are not statistical significant, whereas bigger differences are found for Latin America and sub-Saharan Africa. The results for South Asia and sub-Saharan Africa are of particular interest; it, at least to some extent, explains the difference between South Asia and sub-Saharan Africa with respect to the prevalence of stunting which, is higher in South Asia, while mortality rates are higher in sub-Saharan Africa (Klasen, 2008; Paciorek et al., 2013; You et al., 2015b). The imputed z-scores reveal that the unobserved z-scores of deceased children in sub-Saharan Africa are indeed lower as the observed z-scores. Comparing the imputed z-scores between South Asia and sub-Saharan Africa shows that the discrepancy between the undernutrition rates decreases. Observed z-scores are lowest in Asia and sub-Saharan Africa and male children show lower values as their female counterparts. Additionally, it can be observed that in all three geopolitical regions the nutritional situation measured as height-for-age z-score improved over time.

[Please insert FIGURE 2 about here]

[Please insert FIGURE 3 about here]

Results for country sub-samples. In addition to the main analysis illustrated in the previous section, the effect of selective mortality on undernutrition is analysed at a country level for the 35 countries included in the analysis in order to improve our understanding of selective mortality. Figure 4 and Figure 5 depict the difference between the observed and simulated z-score, Figure B.1 and Figure B.2 in the Appendix show the observed, constructed, and imputed z-scores individually for the 35 countries before and after 2000. We restrict our results to the height-for-age z-score since manifestation of chronic undernutrition is best analysed by stunting, however the results for underweight and wasting are of similar nature.

By looking at the differences between observed z-scores, and constructed z-scores for each of the countries included, the same as the global sample pattern emerges. The differences are negligible due to the aforementioned critique that the biggest share of the constructed z-scores is stemming from the observed z-scores and only around 10% stem from the imputed z-scores of deceased children.

No clear pattern emerges, when comparing the observed height-for-age z-scores with the imputed z-scores of deceased children. Meaning that differences across geopolitical regions vary strongly and structure between geopolitical regions can be seen. Before 2000 around 20%-30% of the included countries show lower and significant imputed z-scores, whereas the vast share of countries show lower but insignificant imputed z-scores and only a minor share of the countries shows higher imputed z-scores compared to the observed z-scores. After 2000 the share of countries having lower imputed z-scores is smaller and generally a tendency to smaller differences between observed z-scores and imputed z-scores can be observed.

Another pattern that arises from the country analysis is that in countries with a strong male preference, such as Bangladesh or India the differences between observed z-score and imputed z-score is higher for girls than for boys. This finding is consistent with the analysis of Bangladesh by Dancer et al. (2008). Our analysis adds a cross-section of several low- and middle-income countries, which adds a broader, more general context. The analysis of Dancer et al. (2008) finds that non-deceased boys born in Bangladesh with an age between one and five years have significantly higher height-for-age z-scores compared to girls. In contrast of this, we find for countries, with a clear and strong male preference like Bangladesh or India (for instance, Fuse, 2010; Khera et al., 2014), that the differences between observed z-scores for height-for-age (as well as for weight-for-age and weight-for-height) to be higher for girls. Within sub-Saharan Africa this pattern was not observed, which is not surprising; as a weak gender preference for boys exists only in a few sub-Saharan African countries (Khera et al., 2014; Rossi & Rouanet, 2015).

[Please insert FIGURE 4 about here]

[Please insert FIGURE 5 about here]

5.2. Modelling the anthropometric status

For an overview of the results using out-of-sample prediction see Appendix C, where the results for the imputed z-scores of deceased children and the overall z-scores are depicted for boys and girls in the two periods.

Results for this approach are comparable to the afore illustrated results of the matching approach, however, the differences between the observed z-scores, and the imputed z-scores for deceased children tend to be larger. This, however, does not change the overall z-scores significantly, due to the small share of deceased children. In line with earlier papers, such as Boerma et al. (1992), who find that overall anthropometric measures are only affected in case the mortality rate is above 50 deaths per 1,000 live births, our results suggest that overall anthropometric measures are only altered in case of even higher mortality rates. In more detail, the imputed z-scores for deceased children pooling the three geopolitical regions are between 6.0% and 9.0% lower compared to the observed z-scores (6.24% pre-2000 boys, 7.69% post-2000 boys; 8.34% pre-2000 girls, 9.39% post-2000 girls). However due to selective mortality, these differences, result only in small changes in the overall z-scores due to selective mortality. The difference between the observed and imputed z-scores are similar when comparing between boys and girls; with boys tending to have smaller differences between the observed and imputed z-scores. Most surprisingly the differences between observed and imputed z-scores are smaller in sub-Saharan Africa (3.54% pre-2000 boys, 4.22% post-2000 boys; 5.51% pre-2000 girls, 6.34% post-2000 girls) than in South Asia (6.44% pre-2000 boys, 10.21% post-2000 boys; 6.73% pre-2000 girls, 11.15% post-2000 girls).

6. COUNTRY CASE STUDY & DISCUSSION

Case study. In this analysis, we estimate the difference between the observed nutritional status of non-deceased children below 36 months and the unobserved nutritional status of deceased children under the counterfactual assumption that these children would have been alive at the day of the interview. This study reveals that the z-score of deceased children, depending on the gender, the period, and the geopolitical region, is between 0.075 and 0.145 standard deviations below the observed z-score of non-deceased children. In addition, while the observed z-score increased on average over time, the difference between the observed and unobserved z-scores decreased, showing to some extent improvements towards achieving the Sustainable Development Goal (SDG) 2[3].

Illustrating three representative examples at a country level (Ghana, India, and Nicaragua) should highlight the new insights, which stem from this analysis. These three countries were chosen for illustration: in the case of Ghana, the country belongs to the group of countries in which the nutritional status over time is stagnating, or even worsening (Van de Poel et al., 2007; Paciorek et al., 2013). In the case of India, the country belongs to the group with a steady improvement for stunting, wasting, and underweight. The prevalence of undernutrition is still found to be high (Paciorek et al., 2013). Additionally, as far as we know only one study exists (Sachdeva et al., 2016), which was conducted with hospitalized children between six and 59 months and for which anthropometric indicators of deceased children is observed. In the case of Nicaragua, the country experienced steady improvements of the nutritional status, with decreasing prevalence of undernutrition (Paciorek et al., 2013).

Figure 4 and Figure 5 depict, amongst others, the differences between observed height-for-age z-scores and imputed z-scores for the three countries by gender. In all three countries the prevalence of stunting is amongst the highest in low- and middle-income countries. The left-hand panel illustrates the results for boys, whereas in the right-hand panel the results for girls are illustrated.

First, countries evolve differently over time: India, despite a still high prevalence of undernutrition, experienced an improvement over time. In Ghana, despite having a better nutritional status than India on average, no improvements of the nutritional status was observed. Nicaragua showed - as the majority of sampled Latin American countries - an improvement in the prevalence of undernutrition, and in particular the stunting rates. These findings are also in line with the current literature (for instance, Paciorek et al., 2013). In addition, the global trend of improved undernutrition is also reflected by the observed increased height-for-age z-scores for both boys, as well as girls in the corresponding time between 1991 and 2015. Although, this was also highlighted by the aforementioned study, this does not take away from the fact that differences between geopolitical regions and countries exist.

Second, the differences between observed and imputed stunting rates for both boys, and girls are bigger in Ghana in contrast to India; with one prominent exemption for India after 2000, where the difference between observed and imputed stunting rates is higher for girls. This anomaly can

---

[3] SDG 2 states: ‘End hunger, achieve food security and improved nutrition, and promote sustainable agriculture’, see `https://sustainabledevelopment.un.org/` for more details about the SDGs.

be interpreted as a result of the strong preference for boys in India Fuse (2010); Khera et al. (2014). Due to, the household resources are probably not evenly distributed between boys and girls, potentially explaining the staggering difference between the differences between observed and imputed stunting rates between genders. This effect of a differing resource allocation was also suggested by Jayachandran & Pande (2017). The more general pattern is that the discrepancy between observed and imputed stunting rates between South Asia and sub-Saharan Africa should decrease, after accounting for selective mortality. This is also due to the rapid progress in South Asian countries to improve the nutritional status of children, whereas this is not observed to the same extent in sub-Saharan Africa (Paciorek et al., 2013).

Third, in general anthropometric indicators improved over time, in particular the height-for-age z-scores (Paciorek et al., 2013). In addition, the difference between the observed and imputed anthropometric indicators decreased over time, and should continue to decrease in the future. This indicates to some extent, that the unobserved anthropometric indicators of deceased children should, on average, also have improved. This similar trend like for the observed anthropometric indicators can be assumed.

Fourth, in the aforementioned study by Sachdeva et al. (2016), anthropometric indicators of hospitalized children between six and 59 months are observed for non-deceased, as well as deceased children. No distinction has been made between boys and girls. The data for this study comes from a hospital mainly treating the poor urban population of East Delhi and urban people of western Uttar Pradesh. As the age of the included children differs from our analysis, no quantitative comparison can be made. Qualitatively the size of the difference between observed and imputed anthropometric indicators corresponds to the numbers of Sachdeva et al. (2016). This strengthens our findings that deceased children show anthropometric indicators between 0.08 and 0.15 standard deviations below the observed anthropometric indicators.

[Please insert FIGURE 6 about here]

Discussion. In this paper the effect of selective mortality on anthropometric indicators is estimated, which explains, to some extent, the so-called South Asian enigma (Ramalingaswami et al., 1996). The phenomena that higher proportions of malnourished and stunted children are observed more persistent in South Asian countries compared to other geopolitical regions. We find first, that before 2000 the proportion of stunted and malnourished children was indeed higher in South Asia, compared to other geopolitical regions. This observation tends to diminish post 2000, where the proportion of malnourished, and stunted children is almost equally high in South Asia and sub-Saharan Africa (compare Table III and Paciorek et al., 2013). Due to this, we assumed the effect of selective mortality to be more prominent in the pre-2000 era.

Using data from six waves of DHS for a sample of 35 low- and middle-income countries, we estimate the effect of selective mortality on the anthropometric indicators of deceased children. Using a similar approach established by Alderman et al. (2011), we are able to expand their results by using a sample of 35 low- and middle-income countries from three geopolitical regions, and find that the results with respect to the 'simulated' overall z-score remain similar. A comparison between the observed anthropometric indicators of non-deceased children and the imputed anthropometric indicators for deceased children adds more insight into the relationship of undernutrition and the risk of dying. We regard the 'simulated' overall anthropometric indicator as potentially misleading. As with declining mortality rates the contribution of the imputed values of deceased children to the overall 'simulated' anthropometric indicators tend to diminish. Additionally, we find no contradicting elements to the analysis conducted by Alderman et al. (2011), even when carrying out the analysis with our sample of 35 low- and middle-income countries. It is important to note that there exists a significant difference between observed anthropometric indicators and the imputed anthropometric indicators of deceased children. We find that anthropometric indicators of deceased children, depending on gender, period, and the geopolitical region are on average between 0.08 and 0.15 standard deviations lower than the observed anthropometric indicators.

In this sense, a comparison between the observed anthropometric indicators and the imputed anthropometric indicators is more likely to contribute to the discussion of the impact of undernutrition on mortality. Due to the current declining mortality rates (for instance, You et al., 2015a), the improvement of the nutritional situation of children (Paciorek et al., 2013) and with fewer children being stunted, the difference between the observed anthropometric indicators and the 'true' value of all children appears to diminish.

Assessing the size of anthropometric indicators for deceased children for a sample of 35 low- and middle-income countries is of particular interest. Since recently a series of papers used multilevel models to estimate the impact of individual, regional/community level, and country-level determinants on mortality (Boco, 2010; Harttgen et al., 2015, 2016, 2017). In these papers, the anthropometric measures are approximated by regional aggregates of non-deceased children, which introduces a bias as the true nutritional status of all children is overestimated. First, we imputed values found in the literature (e.g Pelletier et al., 1995) and used by the WHO as threshold defining the cut-off of outliers. This reveals that selective mortality will only have a significant impact on the overall anthropometric indicators in case the differences between the observed anthropometric indicator of non-deceased children and the unobserved indicator of deceased children is large. Using a survival model and a matching approach, the anthropometric indicators for deceased children are simulated for the counterfactual situation, under which the deceased children were assumed to be alive at the day of data collection. It reveals that deceased children show indeed lower values for the anthropometric indicators. The effects of the nutritional situation in this type of analysis should be seen as lower bound for the effect of undernutrition on mortality. In consideration of this caveat, we expect the 'true' effect of undernutrition on mortality to be of bigger magnitude as found by this multilevel approach.

References

Adedini, S. A., Odimegwu, C., Imasiku, E. N. S., Ononokpono, D. N., & Ibisomi, L. (2015). Regional variations in infant and child mortality in Nigeria: a multilevel analysis. *Journal of Biosocial Science*, 47(2):165–187.

Alderman, H., Lokshin, M., & Radyakin, S. (2011). Tall claims: Mortality selection and the height of children in India. *Economics & Human Biology*, 9(4):393–406.

Arnold, F. (1997). *Gender preferences for children. DHS Comparative Studies No. 23.* Macro International Inc., Calverton, Maryland, USA.

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., Mathers, C., & Rivera, J. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet*, 371(9608):243–260.

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., & Uauy, R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, 382(9890):427–451.

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.

Boerma, J. T., Sommerfelt, A. E., & Bicego, G. T. (1992). Child Anthropometry in Cross-sectional Surveys in Developing Countries: An Assessment of the Survivor Bias. *American Journal of Epidemiology*, 135(4):438–449.

Bryce, J., Boschi-Pinto, C., Shibuya, K., & Black, R. E. (2005). {WHO} estimates of the causes of death in children. *The Lancet*, 365(9465):1147–1152.

Claeson, M., Bos, E. R., Mawji, T., & Pathmanathan, I. (2000). Reducing child mortality in India in the new millennium. *Bulletin of the World Health Organization*, 78:1192–1199.

Cleland, J. G. & 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.

Dancer, D., Rammohan, A., & Smith, M. D. (2008). Infant mortality and child nutrition in Bangladesh. *Health Economics*, 17(9):1015–1035.

Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). *Regression: Models, Methods and Applications.* Springer-Verlag Berlin Heidelberg, 1 edition.

Fenske, N., Burns, J., Hothorn, T., & Rehfuess, E. A. (2013). Understanding Child Stunting in India: A Comprehensive Analysis of Socio-Economic, Nutritional and Environmental Determinants Using Additive Quantile Regression. *PLOS ONE*, 8(11):1–16.

Fuse, K. (2010). Variations in attitudinal gender preferences for children across 50 less-developed countries. *Demographic Research*, 23(36):1031–1048.

Gørgens, T., Meng, X., & Vaithianathan, R. (2012). Stunting and selection effects of famine: A case study of the Great Chinese Famine. *Journal of Development Economics*, 97(1):99–111.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., & Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. *The Lancet*, 369(9555):60–70.

Gutierrez, R. G. (2002). Parametric frailty and shared frailty survival models. *Stata Journal*, 2(1):22–44.

Harttgen, K. & Günther, I. (2011). Deadly Cities? A Note on Spatial Inequalities in Mortality in Sub-Saharan Africa. Technical Report 39, ETH Zurich, Kiel und Hamburg.

Harttgen, K., Lang, S., & Santer, J. (2015). Multilevel modelling of child mortality in Africa. Working Paper 03, Working Papers in Economics and Statistics University of Innsbruck.

Harttgen, K., Lang, S., & Santer, J. (2016). A multilevel approach to estimate non-linear effects in determinants of under-5 mortality in Sub-Saharan Africa. Technical report, University of Innsbruck.

Harttgen, K., Santer, J., Lang, S., & Seiler, J. (2017). Modeling Under-5 Mortality through Multilevel Structured Additive Regression with Varying Coefficients for Asia and Sub-Saharan Africa. Working Paper 15, Working Papers in Economics and Statistics University of Innsbruck.

Hastie, T. & Tibshirani, R. (1993). Varying-Coefficient Models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(4):757–796.

Heckman, J. J., Lalonde, R. J., & Smith, J. A. (1999). Chapter 31 - The Economics and Econometrics of Active Labor Market Programs. In Ashenfelter, O. C. & Card, D., editors, *Handbook of Labor Economics*, volume 3, Part A of *Handbook of Labor Economics*, pages 1865–2097. Elsevier.

Jayachandran, S. & Pande, R. (2017). Why Are Indian Children So Short? The Role of Birth Order and Son Preference. *American Economic Review*, 107(9):2600–2629.

Jorde, L. B. & Wooding, S. P. (2004). Genetic variation, classification and 'race'. *Nature Genetics*, 36:S28–33.

Khera, R., Jain, S., Lodha, R., & Ramakrishnan, S. (2014). Gender bias in child care and child health: global patterns. *Archives of Disease in Childhood*, 99(4):369–374.

Klasen, S. (2008). Poverty, undernutrition, and child mortality: Some inter-regional puzzles and their implicationsfor research and policy. *The Journal of Economic Inequality*, 6(1):89–115.

Lang, S., Umlauf, N., Wechselberger, P., Harttgen, K., & Kneib, T. (2014). Multilevel structured additive regression. *Statistics and Computing*, 24(2):223–238.

Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., Mathers, C., & Black, R. E. (2012). Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *The Lancet*, 379(9832):2151–2161.

Liu, L., Oza, S., Hogan, D., Perin, J., Rudan, I., Lawn, J. E., Cousens, S., Mathers, C., & Black, R. E. (2015). Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. *The Lancet*, 385(9966):430–440.

Martorell, R. & Habicht, J.-P. (1986). Growth in early childhood in developing countries. In Falkner, F. T. & Tanner, J. M., editors, *Human growth: A Comprehensive Treatise*, volume 3, chapter 12, pages 241–62. Plenum Press, New York, 2 edition.

Misselhorn, M. & 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.

Moradi, A. (2010). Selective mortality or growth after childhood? What really is key to understand the puzzlingly tall adult heights in sub-Saharan Africa. Technical report, Centre for the Study of African Economies, University of Oxford.

Mosley, W. H. & Chen, L. C. (1984). An Analytical Framework for the Study of Child Survival in Developing Countries. *Population and Development Review*, 10:25–45.

Munda, M., Rotolo, F., & Legrand, C. (2012). parfm: Parametric Frailty Models in R. *Journal of Statistical Software*, 51(1):1–20.

Paciorek, C. J., Stevens, G. A., Finucane, M. M., & Ezzati, M. (2013). Children's height and weight in rural and urban populations in low-income and middle-income countries: a systematic analysis of population-representative data. *The Lancet Global Health*, 1(5):e300–e309.

Pelletier, D. L. & 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.

Pelletier, D. L., Frongillo, E. A., Schroeder, D. G., & Habicht, J. P. (1995). The effects of malnutrition on child mortality in developing countries. *Bulletin of the World Health Organization*, 73(4):443–448.

Pitt, M. M. (1997). Estimating the Determinants of Child Health When Fertility and Mortality Are Selective. *The Journal of Human Resources*, 32(1):129–158.

Ramalingaswami, V., Jonsson, U., & Rohde, J. (1996). The Asian enigma. In *The progress of nations*. New York New York UNICEF 1996.

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rossi, P. & Rouanet, L. (2015). Gender Preferences in Africa: A Comparative Analysis of Fertility Choices. *World Development*, 72(Supplement C):326–345.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.

Rubin, D. B. (1979). Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association*, 74(366):318–328.

Sachdeva, S., Dewan, P., Shah, D., Malhotra, R. K., & Gupta, P. (2016). Mid-upper arm circumference v. weight-for-height Z-score for predicting mortality in hospitalized children under 5 years of age. *Public Health Nutrition*, 19(14):2513—-2520.

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

UNICEF (1990). Strategy for Improved Nutrition of Children and Women in Developing Countries. A UNICEF Policy Review.

Van de Poel, E., Hosseinpoor, A. R., Jehu-Appiah, C., Vega, J., & Speybroeck, N. (2007). Malnutrition and the disproportional burden on the poor: the case of Ghana. *International Journal for Equity in Health*, 6(1):21.

Vollmer, S., Bommer, C., Krishna, A., Harttgen, K., & Subramanian, S. (2017). The association of parental education with childhood undernutrition in low- and middle-income countries: comparing the role of paternal and maternal education. *International Journal of Epidemiology*, 46(1):312–323.

WHO (2010). *WHO Anthro for personal computers, version 3.2.2, 2011: Software for assessing growth and development of the world's children*. WHO, Geneva.

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

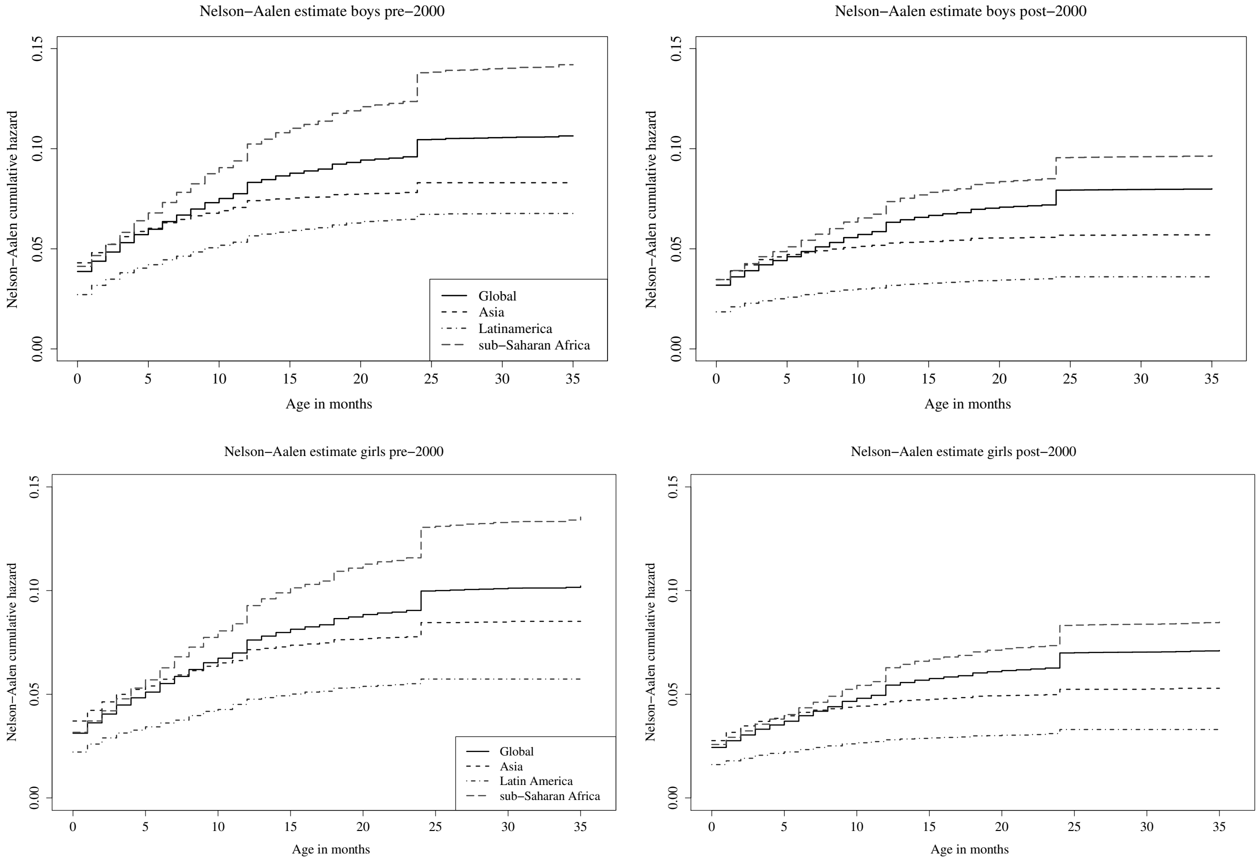
WHO Multicentre Growth Reference Study Group & de Onis, M. (2006). Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. *Acta Paediatrica*, 95(S450):56–65.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1):3–36.

You, D., Hug, L., Ejdemyr, S., Idele, P., Hogan, D., Mathers, C., Gerland, P., New, J., & Alkema, L. (2015a). 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.

You, D., New, J., & Wardlaw, T. (2015b). Levels and trends in child mortality. Report 2015. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation. Technical report, United Nations.

7. Figures



Nelson–Aalen estimate boys pre–2000

Nelson–Aalen estimate boys post–2000

Nelson–Aalen estimate girls pre–2000

Nelson–Aalen estimate girls post–2000

Figure 1: Nelson-Aalen cumulative hazard by gender, geopolitical region, and before and after 2000. The Nelson-Aalen cumulative hazard for boys and girls are shown in the top and bottom panels, respectively. **Source:** Own calculation with available data from DHS survey.

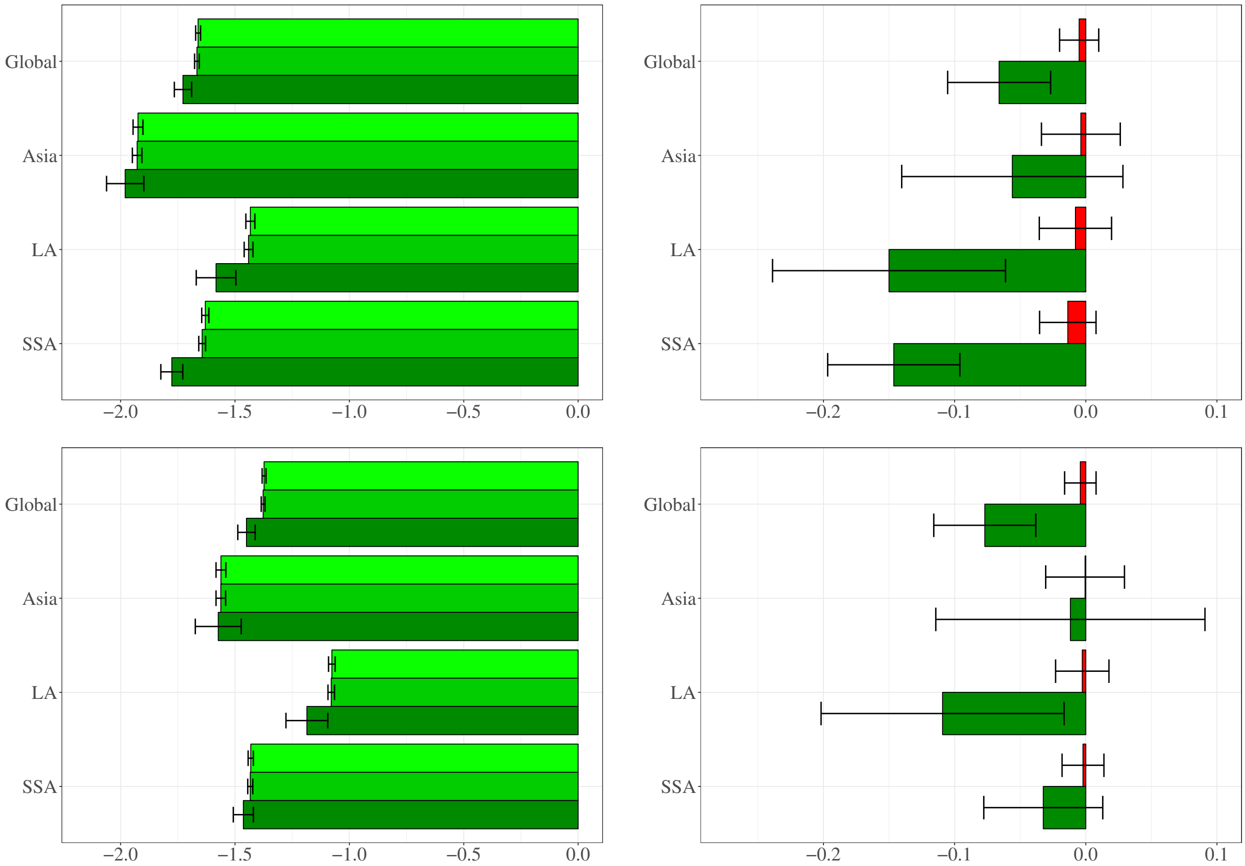


Figure 2: Selective mortality and its impact on stunting for boys. The left-hand panels depict the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. The right-hand panels depict the difference between the observed and the simulated z-scores (**grey**) and between the observed and the imputed z-scores (**dark grey**) by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. **Source:** Own calculation with available data from DHS survey.

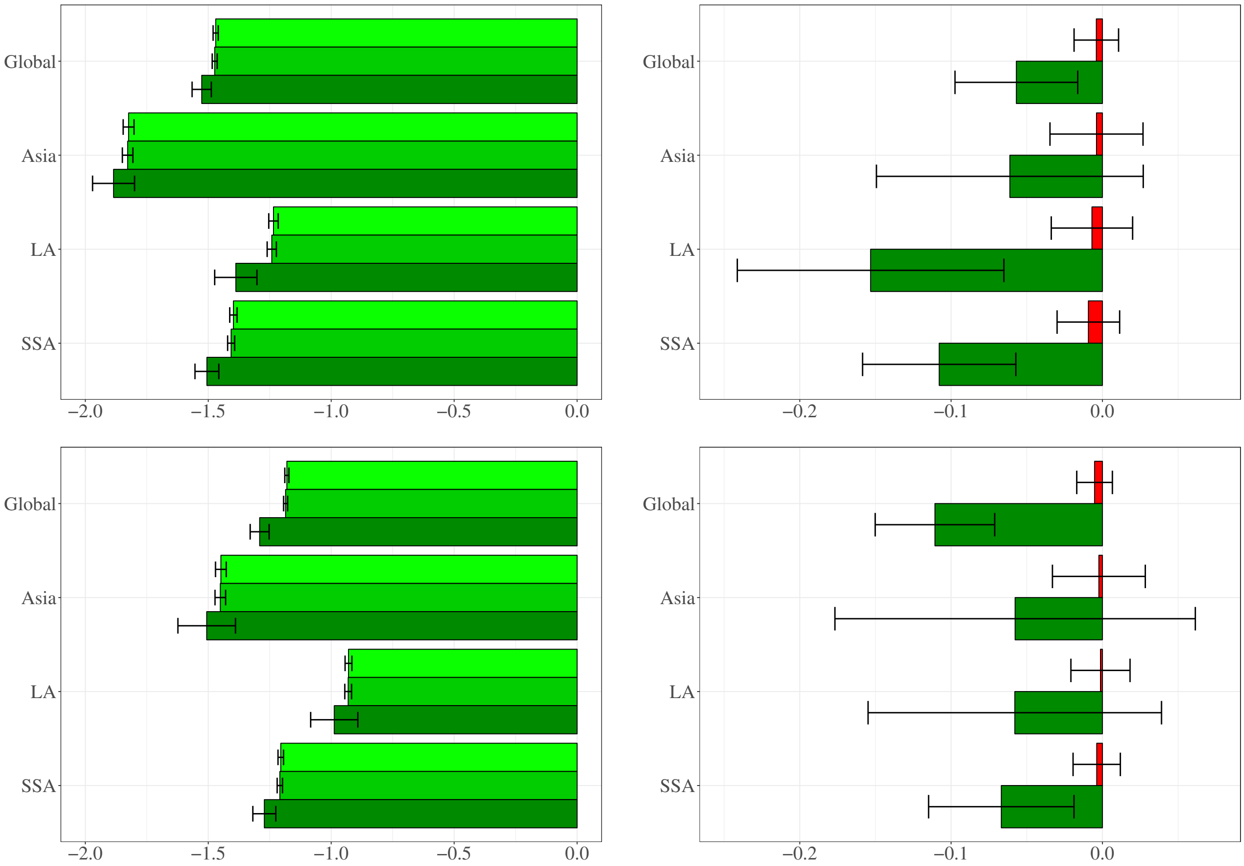


Figure 3: Selective mortality and its impact on stunting for girls. The left-hand panels depict the observed (light grey), simulated (grey), and imputed (**dark grey**) z-scores by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. The right-hand panels depict the difference between the observed and the simulated z-scores (grey) and between the observed and the imputed z-scores (**dark grey**) by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. **Source:** Own calculation with available data from DHS survey.

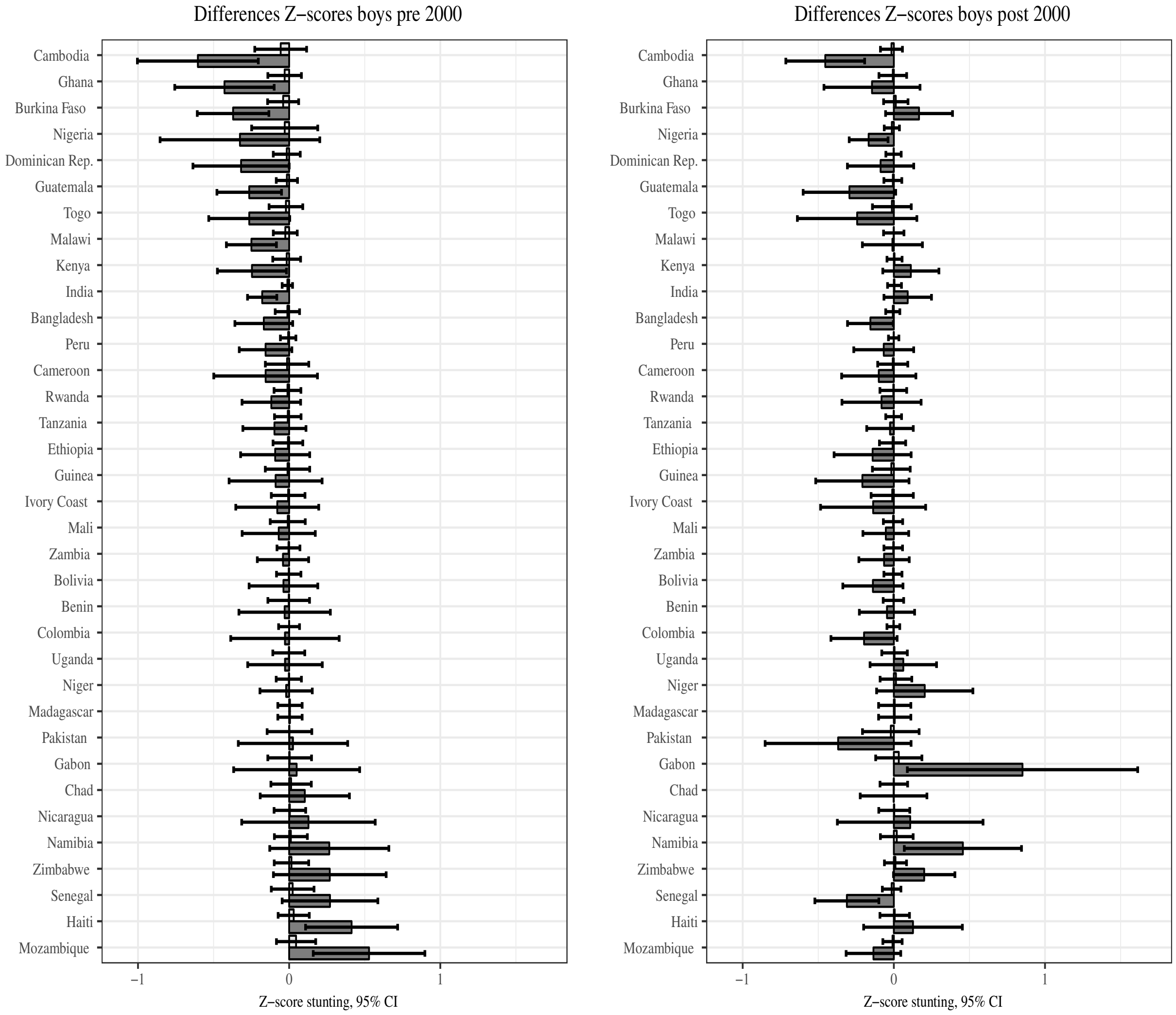


Figure 4: Selective mortality and its impact on stunting for boys by country. The left-hand panels depict the difference between the observed and the simulated z-scores (**grey**) and between the observed and the imputed z-scores (**dark grey**) by country, together with 95% confidence intervals for the period before 2000. The right-hand panels depict the difference between the observed and the simulated z-scores (**grey**) and between the observed and the imputed z-scores (**dark grey**) by country, together with 95% confidence intervals for the period after 2000. **Source:** Own calculation with available data from DHS survey.

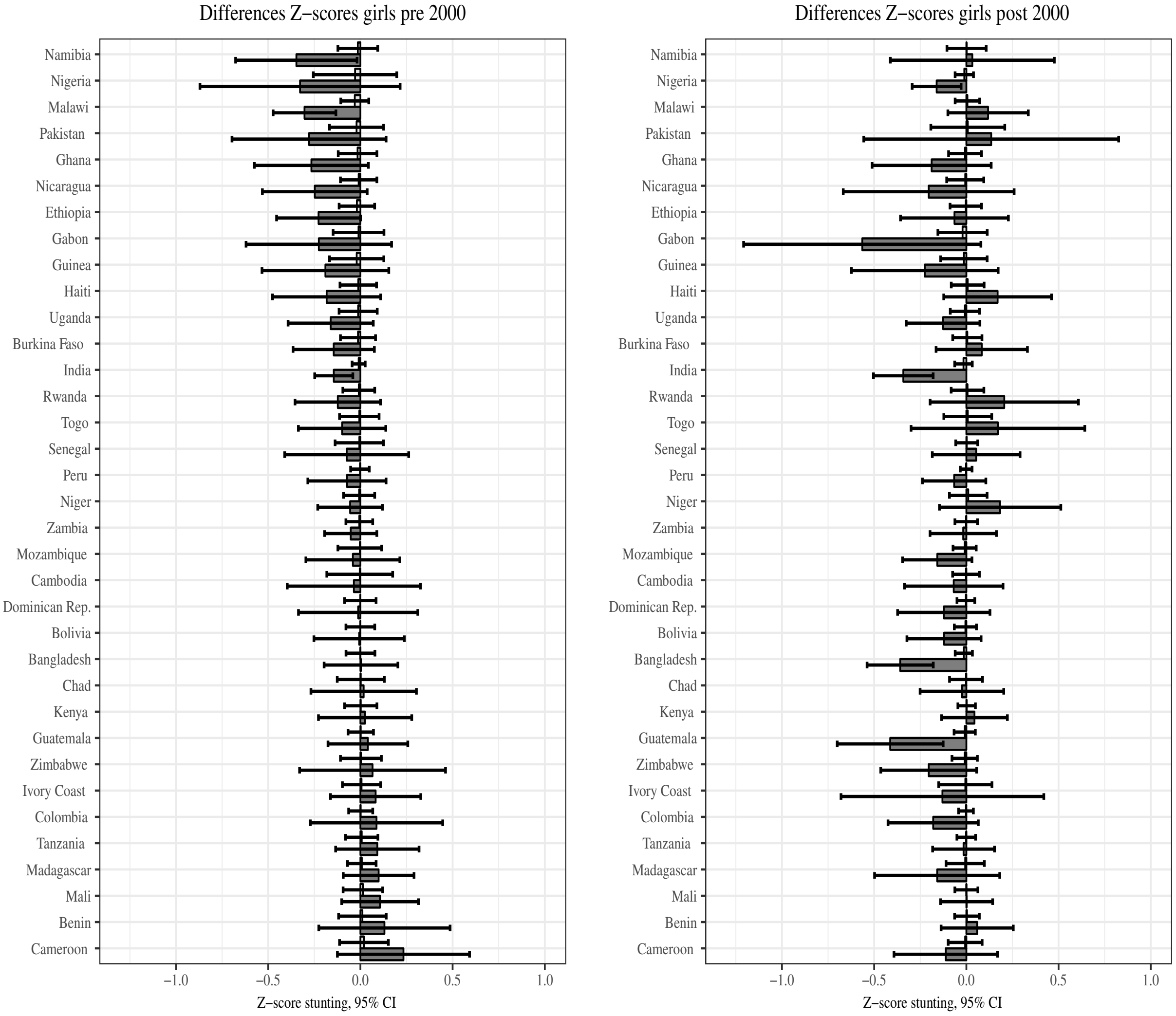


Figure 5: Selective mortality and its impact on stunting for girls by country. The left-hand panels depict the difference between the observed and the simulated z-scores (**grey**) and between the observed and the imputed z-scores (**dark grey**) by country, together with 95% confidence intervals for the period before 2000. The right-hand panels depict the difference between the observed and the simulated z-scores (**grey**) and between the observed and the imputed z-scores (**dark grey**) by country, together with 95% confidence intervals for the period after 2000. **Source:** Own calculation with available data from DHS survey.

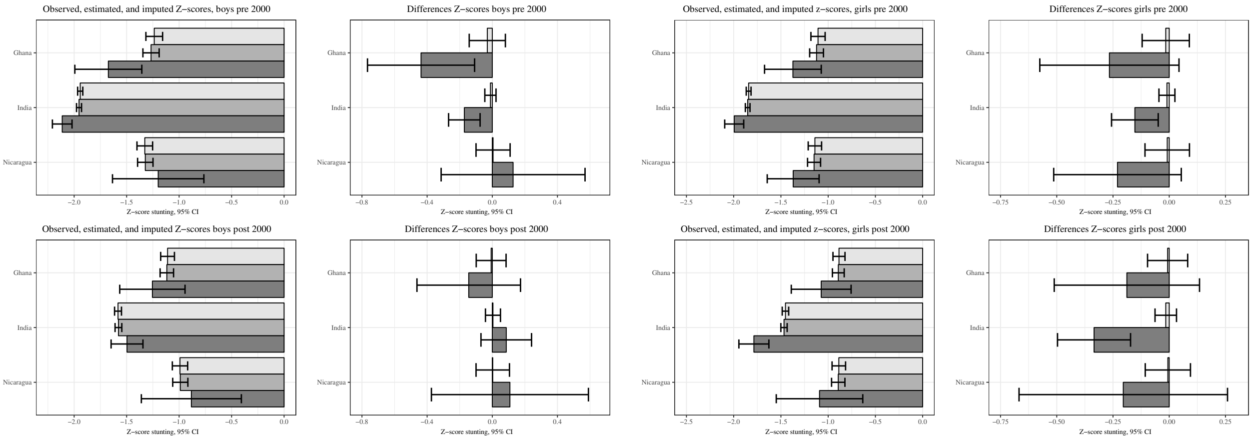


Figure 6: Selective mortality and its impact on stunting for selected countries. The left-hand panels depict the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores for boys, together with 95% confidence intervals for the periods before and after 2000. The right-hand panels depict the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores for boys, together with 95% confidence intervals for the periods before and after 2000. **Source:** Own calculation with available data from DHS survey.

8. Tables

Table I: Share of deaths by period, region, age, and gender (in %).

Period	**Age**	**Global**	**Asia**	**LA**	**SSA**	**Global**	**Asia**	**LA**	**SSA**
Pre 2000	Neonatal	47.064	60.417	48.464	40.451	40.978	53.485	46.728	33.736
	01-06 months	26.368	24.591	27.826	26.822	27.522	25.464	26.743	28.675
	07-12 months	15.929	10.305	14.957	18.813	18.636	13.73	17.212	21.283
	13-18 months	5.585	2.087	5.159	7.333	6.868	3.468	6.046	8.665
	19-24 months	4.715	2.6	3.42	6.041	5.578	3.713	3.272	6.995
	25-35 months	0.339	0	0.174	0.541	0.417	0.14	0	0.645
Post 2000	Neonatal	50.321	66.667	58.418	47.108	44.711	59.961	56.926	41.131
	01-06 months	23.526	23.289	24.746	23.428	24.945	26.627	23.311	24.886
	07-12 months	15.793	6.567	10.522	17.683	18.609	7.758	12.247	20.894
	13-18 months	5.34	2.373	4.064	5.902	5.694	3.09	3.716	6.293
	19-24 months	4.868	1.049	2.25	5.697	5.769	2.433	3.801	6.473
	25-35 months	0.151	0.055	0	0.181	0.272	0.131	0	0.323

Source: DHS data; calculation by authors.

Table II: Descriptive statistic anthropometric indicator (stunting) by period, region, and gender.

Period		**Global**	**Asia**	**LA**	**SSA**	**Global**	**Asia**	**LA**	**SSA**
Pre 2000	Mean	-1.661	-1.924	-1.433	-1.63	-1.47	-1.824	-1.235	-1.398
	SD	1.813	1.884	1.63	1.84	1.764	1.881	1.571	1.762
	n	108,980	29,250	26,259	53,471	106,922	27,875	25,652	53,395
Post 2000	Mean	-1.373	-1.561	-1.077	-1.431	-1.18	-1.449	-0.93	-1.205
	SD	1.824	1.735	1.434	1.947	1.776	1.727	1.355	1.895
	n	173,412	25,080	37,713	110,619	170,522	23,957	36,563	110,002

Source: DHS data; calculation by authors.

Table III: Share of non-, moderately-, and severely stunted observations by period, region, and gender.

Period	**Stunted**	**Global**	**Asia**	**LA**	**SSA**	**Global**	**Asia**	**LA**	**SSA**
Pre 2000	non	0.563	0.495	0.64	0.563	0.616	0.522	0.696	0.627
	moderately	0.213	0.222	0.201	0.215	0.203	0.217	0.183	0.205
	severely	0.223	0.283	0.159	0.222	0.181	0.261	0.122	0.168
Post 2000	not	0.635	0.59	0.753	0.605	0.688	0.619	0.799	0.666
	moderately	0.193	0.222	0.165	0.196	0.176	0.219	0.145	0.177
	severely	0.172	0.189	0.083	0.199	0.136	0.162	0.056	0.157

According to the WHO children are not stunted if the z-score is $<$ -2; moderately if the z-score is $\leq$ -2 and $<$ -3; and severely if the z-score is $\leq$ -3. **Source:** DHS data; calculation by authors.

Table IV: Changes in the height-for-age z-score (Z_{all}) for different imputation scenarios.

Period		**Region**	**Z_d boys**	**Z_s boys**	**Z_{all} boys**	**Δ**	**Z_d girls**	**Z_s girls**	**Z_{all} girls**	**Δ**
Pre 2000	I1	**Global**	-1.771	-1.661	-1.673	-0.012	-1.582	-1.47	-1.481	-0.011
		Asia	-2.206	-1.924	-1.947	-0.023	-2.106	-1.824	-1.844	-0.02
		LA	-1.749	-1.433	-1.466	-0.034	-1.599	-1.235	-1.272	-0.037
		SSA	-1.749	-1.63	-1.639	-0.01	-1.517	-1.398	-1.406	-0.009
	I2	**Global**	-3.323	-1.661	-1.838	-0.177	-2.94	-1.47	-1.62	-0.15
		Asia	-3.848	-1.924	-2.078	-0.154	-3.648	-1.824	-1.954	-0.13
		LA	-2.866	-1.433	-1.585	-0.152	-2.47	-1.235	-1.361	-0.126
		SSA	-3.26	-1.63	-1.76	-0.13	-2.795	-1.398	-1.497	-0.099
	I3	**Global**	-6	-1.661	-2.123	-0.462	-6	-1.47	-1.933	-0.463
		Asia	-6	-1.924	-2.25	-0.326	-6	-1.824	-2.121	-0.297
		LA	-6	-1.433	-1.919	-0.486	-6	-1.235	-1.722	-0.487
		SSA	-6	-1.63	-1.979	-0.35	-6	-1.398	-1.725	-0.327
	I4	**Global**	-2.525	-1.661	-1.753	-0.092	-2.461	-1.47	-1.571	-0.101
		Asia	-2.611	-1.561	-1.645	-0.084	-2.578	-1.449	-1.529	-0.08
		LA	-2.33	-1.077	-1.21	-0.133	-2.282	-0.93	-1.068	-0.138
		SSA	-2.514	-1.63	-1.701	-0.071	-2.438	-1.398	-1.472	-0.074
Post 2000	I1	**Global**	-1.46	-1.373	-1.382	-0.009	-1.27	-1.18	-1.189	-0.009
		Asia	-1.842	-1.561	-1.584	-0.022	-1.728	-1.449	-1.468	-0.02
		LA	-1.309	-1.077	-1.101	-0.025	-1.193	-0.93	-0.957	-0.027
		SSA	-1.519	-1.431	-1.438	-0.007	-1.295	-1.205	-1.212	-0.006
	I2	**Global**	-2.746	-1.373	-1.519	-0.146	-2.361	-1.18	-1.301	-0.121
		Asia	-3.123	-1.561	-1.686	-0.125	-2.897	-1.449	-1.552	-0.103
		LA	-2.153	-1.077	-1.191	-0.115	-1.859	-0.93	-1.025	-0.095
		SSA	-2.862	-1.431	-1.546	-0.114	-2.41	-1.205	-1.291	-0.086
	I3	**Global**	-6	-1.373	-1.865	-0.492	-6	-1.18	-1.673	-0.493
		Asia	-6	-1.561	-1.916	-0.355	-6	-1.449	-1.772	-0.323
		LA	-6	-1.077	-1.6	-0.524	-6	-0.93	-1.448	-0.518
		SSA	-6	-1.431	-1.797	-0.365	-6	-1.205	-1.546	-0.341
	I4	**Global**	-2.429	-1.373	-1.485	-0.112	-2.366	-1.18	-1.301	-0.121
		Asia	-2.492	-1.561	-1.636	-0.074	-2.454	-1.449	-1.52	-0.071
		LA	-2.33	-1.077	-1.21	-0.133	-2.282	-0.93	-1.068	-0.138
		SSA	-2.449	-1.431	-1.512	-0.081	-2.374	-1.205	-1.288	-0.083

I1: Imputed hazard (Z_d) that results in a statistically significant change in the total hazard (Z_{all}) at $\alpha = 0.05$. I2: $Z_d = 2 \times Z_s$. I3: Lower bounds of WHO considered as outliers. I4: Upper bounds found by Pelletier et al. (1995). **Source:** DHS data; calculation by authors.

Table V: Descriptive statistics main explanatory variables

Period	**Variable**	**Mean, %**	**SD**	**Mean, %**	**SD**	**Mean, %**	**SD**	**Mean, %**	**SD**
Pre 2000	Male	50.828%		51.529%		50.939%		50.317%	
	Age child (in months)	15.987	10.582	15.974	10.643	16.807	10.527	15.613	10.546
	First born	24.529%		27.816%		26.633%		21.398%	
	Second born	20.496%		24.293%		20.945%		17.801%	
	Third born	15.675%		17.395%		15.661%		14.557%	
	Fourth born	11.435%		11.192%		10.692%		11.94%	
	Fifth born	8.618%		7.41%		7.687%		9.843%	
	Sixth born	6.5%		4.897%		5.706%		7.919%	
	Seventh born	4.698%		3.063%		4.253%		5.975%	
	Eigth born (or higher)	8.049%		3.935%		8.423%		10.568%	
	Age mother (in years)	27.05	6.497	25.806	5.707	27.627	6.745	27.595	6.748
	Education (in years)	3.889	2.39	3.3	1.639	3.6	1.733	4.414	2.958
	No edu. mother	42.497%		54.471%		18.557%		45.825%	
	Prim. edu. mother	34.971%		18.877%		46.375%		40.186%	
	Sec. edu. mother	18.898%		21.342%		27.423%		13.324%	
	Higher edu. mother	3.633%		5.309%		7.644%		0.666%	
	Household size	6.887	3.359	7.042	3.325	6.338	2.731	7.043	3.61
	Share children under 5	0.28	0.152	0.244	0.156	0.302	0.151	0.293	0.146
	Share adults	0.461	0.175	0.515	0.176	0.455	0.174	0.43	0.167
	Urban	31.401%		27.404%		46.623%		26.919%	
	Asset index	-0.09	0.999	0.003	0.975	-0.363	1.01	-0.023	0.987
	n	275,204		84,927		60,515		129,762	
Post 2000	Male	50.83%		51.38%		50.925%		50.697%	
	Age child (in months)	16.381	10.512	16.755	10.6	17.226	10.422	16.095	10.504
	First born	24.911%		33.149%		33.132%		21.208%	
	Second born	21.063%		27.079%		25.23%		18.819%	
	Third born	15.877%		16.228%		16.707%		15.599%	
	Fourth born	11.725%		9.465%		9.33%		12.776%	
	Fifth born	8.651%		5.736%		5.551%		10.009%	
	Sixth born	6.442%		3.456%		3.721%		7.719%	
	Seventh born	4.511%		2.112%		2.468%		5.502%	
	Eigth born (or higher)	6.82%		2.775%		3.86%		8.369%	
	Age mother (in years)	27.497	6.591	26.227	5.819	27.129	6.747	27.843	6.66
	Education (in years)	4.063	2.398	3.552	1.747	3.915	2.219	4.248	2.577
	No edu. mother	36.464%		32.196%		8.451%		44.34%	
	Prim. edu. mother	35.603%		26.184%		40.927%		36.144%	
	Sec. edu. mother	23.239%		34.172%		37.692%		17.436%	
	Higher edu. mother	4.694%		7.448%		12.93%		2.08%	
	Household size	6.626	3.255	6.668	3.041	5.805	2.485	6.823	3.43
	Share children under 5	0.306	0.141	0.273	0.142	0.3	0.137	0.314	0.14
	Share adults	0.441	0.162	0.517	0.16	0.477	0.164	0.417	0.155
	Urban	32.791%		33.865%		54.26%		27.193%	
	Asset index	-0.14	1.009	0.006	0.999	-0.363	1.165	-0.113	0.959
	n	493,840		67,857		85,406		340,577	

Descriptive statistic includes children younger three years, for whom all covariates included in the Weibull model were complete.

Source: DHS data; calculation by authors.

Table VI: Results Weibull proportional hazard estimation for children younger 36 months.

Variable	**Hazard Ratio**	**SE**	**Hazard Ratio**	**SE**	**Hazard Ratio**	**SE**	**Hazard Ratio**	**SE**
Constant	-3.2394***	0.167	-3.3936***	0.131	-2.9926***	0.211	-2.4102***	0.151
Male child	1.1632***	0.026	1.0567	0.043	1.2985***	0.073	1.1794***	0.036
Second born	1.4622***	0.054	1.196**	0.072	2.5306***	0.240	1.3031***	0.067
Third born	1.9584***	0.086	1.5655***	0.114	3.577***	0.401	1.6564***	0.101
Fourth born	2.4039***	0.123	2.1983***	0.190	5.1734***	0.665	1.7301***	0.122
Fisth born	2.8341***	0.165	2.8239***	0.283	5.6539***	0.829	1.9776***	0.156
Sixth born	3.3342***	0.216	3.071***	0.360	8.3771***	1.349	2.1964***	0.191
Seventh born	4.0522***	0.293	4.8002***	0.648	10.0444***	1.765	2.4266***	0.233
Eighth born (or higher)	5.4502***	0.384	6.2876***	0.845	12.4813***	2.123	3.4138***	0.320
Age Mother	0.5656***	0.012	0.5967***	0.020	0.4961***	0.025	0.6171***	0.018
Age Mother^2	1.116***	0.013	1.1001***	0.018	1.1485***	0.032	1.1008***	0.018
Primary education	0.8106***	0.024	0.7392***	0.041	0.8711·	0.065	0.8375***	0.033
Sec. or higher edu.	0.4903***	0.020	0.5077***	0.031	0.5189***	0.055	0.5238***	0.033
Household size	0.2508***	0.025	0.2271***	0.032	0.3313***	0.061	0.4749**	0.112
Household size^2	0.5876***	0.010	0.6692***	0.018	0.5706***	0.023	0.618***	0.015
Share children 0-5	1.1419***	0.009	1.1034***	0.016	1.1107***	0.021	1.1392***	0.014
Share adults	7e-04***	0.000	0.006***	0.001	8e-04***	0.000	4e-04***	0.000
Second quintile AI	2.8749***	0.226	4.1127***	0.551	1.9626***	0.394	2.7249***	0.293
Third quintile AI	0.9371*	0.031	0.8189***	0.046	0.9876	0.076	1.0212	0.046
Fourth quintile AI	0.8337***	0.029	0.6264***	0.039	0.8408·	0.075	1.0086	0.046
Fifth quintile AI	0.7269***	0.028	0.5725***	0.040	0.7998*	0.084	0.8439***	0.043
Urban	0.6143***	0.030	0.462***	0.040	0.563***	0.072	0.7587***	0.049
Country FE	Yes		Yes		Yes		Yes	
Year FE	Yes		No		No		No	
$\ln(\theta)$	3.9892***	0.064	3.5821***	0.096	3.7652***	0.176	3.9446***	0.084
$\ln(\rho)$	0.2754***	0.010	0.0896***	0.018	0.1182***	0.030	0.3711***	0.013
Log-Likelihood	-105,766.04		-29,746.56		-16,387.54		-59,409.26	
n	275,204		84,927		60,515		129,762	

$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. **Source:** DHS data; calculation by authors.

Table VII: Results Weibull proportional hazard estimation for children younger 36 months.

Variable	**Hazard Ratio**	**SE**	**Hazard Ratio**	**SE**	**Hazard Ratio**	**SE**	**Hazard Ratio**	**SE**
Constant	-4.107***	0.067	-5.1594***	0.062	-3.5048***	0.176	-2.921***	0.092
Male child	1.311***	0.026	1.1409***	0.042	1.1478**	0.058	1.339***	0.030
Second born	2.1427***	0.070	2.1589***	0.117	2.8539***	0.239	1.9594***	0.076
Third born	2.9747***	0.117	3.5956***	0.241	5.015***	0.525	2.4876***	0.114
Fourth born	3.8346***	0.173	5.3186***	0.426	6.3527***	0.808	3.1583***	0.164
Fisth born	4.8867***	0.250	8.8219***	0.812	8.2685***	1.220	3.8593***	0.226
Sixth born	6.285***	0.357	11.528***	1.260	10.2313***	1.708	4.9294***	0.317
Seventh born	7.3143***	0.465	13.4303***	1.754	12.3203***	2.339	5.8282***	0.414
Eighth born (or higher)	11.1445***	0.693	21.1204***	2.671	20.4361***	3.707	8.6372***	0.605
Age Mother	0.5464***	0.010	0.6027***	0.018	0.5319***	0.024	0.5569***	0.012
Age Mother^2	1.1139***	0.011	1.0848***	0.015	1.1047***	0.028	1.1102***	0.013
Primary education	0.8113***	0.022	0.8279***	0.041	0.7287***	0.061	0.8362***	0.025
Sec. or higher edu.	0.5665***	0.020	0.6557***	0.035	0.6233***	0.063	0.5643***	0.024
Household size	0.3656***	0.028	0.4079***	0.050	0.5441***	0.075	0.3789***	0.041
Household size^2	0.4495***	0.007	0.4134***	0.012	0.4511***	0.018	0.5078***	0.010
Share children 0-5	1.235***	0.008	1.2577***	0.014	1.1938***	0.017	1.2026***	0.010
Share adults	1e-04***	0.000	5e-04***	0.000	2e-04***	0.000	1e-04***	0.000
Second quintile AI	3.9722***	0.273	12.1042***	1.318	3.7363***	0.656	3.8638***	0.321
Third quintile AI	0.9182**	0.026	0.9142·	0.045	0.8079**	0.058	0.9855	0.032
Fourth quintile AI	0.8543***	0.026	0.8053***	0.046	0.7906**	0.064	0.9237*	0.031
Fifth quintile AI	0.8047***	0.027	0.7712***	0.051	0.7153***	0.065	0.8908**	0.033
Urban	0.6474***	0.027	0.5847***	0.050	0.6389***	0.066	0.7286***	0.034
Country FE	Yes		Yes		Yes		Yes	
Year FE	Yes		No		No		No	
$\ln(\theta)$	4.2938***	0.064	-0.2383	0.302	2.5519***	0.229	4.258***	0.063
$\ln(\rho)$	0.2537***	0.009	-0.3421***	0.019	-0.1591***	0.032	0.3***	0.009
Log-Likelihood	-145,320.03		-16,341.37		-13,891.31		-114,815.22	
n	493,840		67,857		85,406		340,577	

$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. **Source:** DHS data; calculation by authors.

A. Mortality rates, z-scores, and observations by country and survey year

Table A.1: Under-3 mortality rates per 1,000 live births (U3M), z-score stunting, and number of observations (n), by country, and period.

Country	Year	U3M	Stunting	n	Country	Year	U3M	Stunting	n
Bangladesh	1997	104	-2.078	3,555	Madagascar	1997	127	-2.034	3,643
Bangladesh	2000	87	-1.808	4,086	Madagascar	2004	67	-1.749	3,184
Bangladesh	2004	73	-1.762	4,057	Madagascar	2009	62	-1.450	7,281
Bangladesh	2007	51	-1.551	3,606	Malawi	1992	211	-1.793	2,772
Bangladesh	2011	49	-1.495	4,934	Malawi	2000	166	-1.734	7,433
Bangladesh	2014	41	-1.372	4,696	Malawi	2004	108	-1.842	6,734
Benin	1996	138	-1.301	2,762	Malawi	2010	99	-1.681	11,827
Benin	2001	141	-1.278	3,096	Malawi	2015	54	-1.141	10,098
Benin	2006	101	-1.569	9,588	Mali	1996	208	-1.328	5,849
Benin	2012	63	-1.466	7,758	Mali	2001	200	-1.386	7,791
Bolivia	1994	97	-1.294	3,562	Mali	2006	162	-1.199	8,290
Bolivia	1998	85	-1.332	4,137	Mali	2013	73	-1.106	5,942
Bolivia	2003	65	-1.269	5,817	Mozambique	1997	126	-1.684	3,967
Bolivia	2008	53	-1.057	5,134	Mozambique	2003	136	-1.714	5,886
Burkina Faso	1993	136	-1.109	3,176	Mozambique	2011	77	-1.516	6,840
Burkina Faso	1999	168	-1.453	3,260	Namibia	1992	68	-1.380	2,315
Burkina Faso	2003	147	-1.269	5,783	Namibia	2000	72	-0.946	2,317
Burkina Faso	2010	104	-1.221	8,818	Namibia	2007	70	-1.145	3,079
Cambodia	2000	117	-1.481	4,618	Namibia	2013	49	-0.794	3,008
Cambodia	2005	73	-1.638	4,933	Nicaragua	1998	55	-1.233	4,659
Cambodia	2010	50	-1.455	4,836	Nicaragua	2001	44	-0.940	4,105
Cambodia	2014	25	-1.243	4,357	Niger	1992	249	-1.473	4,009
Cameroon	1991	128	-1.357	712	Niger	1998	187	-1.781	4,496
Cameroon	1998	132	-1.248	2,296	Niger	2006	147	-1.706	5,331
Cameroon	2004	136	-1.268	4,735	Niger	2012	95	-1.424	7,338
Cameroon	2011	104	-1.060	6,972	Nigeria	1999	165	-1.886	3,280
Chad	1997	180	-1.333	4,308	Nigeria	2003	177	-1.443	3,592
Chad	2004	176	-1.260	3,181	Nigeria	2008	156	-1.349	16,509
Chad	2015	124	-1.265	9,991	Nigeria	2013	106	-1.058	18,476
Colombia	1995	32	-0.926	3,039	Pakistan	1991	97	-1.928	3,801
Colombia	2000	23	-0.936	2,799	Pakistan	2013	71	-1.565	6,536
Colombia	2005	25	-0.790	8,571	Peru	1992	77	-1.357	5,313
Colombia	2010	19	-0.786	10,305	Peru	1996	57	-1.264	9,210
Cote d'Ivoire	1994	131	-1.241	3,542	Peru	2000	39	-1.288	7,428
Cote d'Ivoire	1999	172	-1.009	1,136	Peru	2009	25	-1.209	5,339
Cote d'Ivoire	2012	99	-1.058	4,599	Peru	2010	23	-1.216	4,983
Dominican Republic	1991	71	-1.116	2,615	Peru	2011	18	-1.180	5,010
Dominican Republic	1996	46	-0.694	2,720	Peru	2012	20	-1.111	5,214
Dominican Republic	2002	34	-0.545	6,830	Rwanda	1992	138	-1.825	3,265
Dominican Republic	2007	37	-0.471	6,431	Rwanda	2000	153	-1.414	4,517
Dominican Republic	2013	31	-0.365	2,193	Rwanda	2005	113	-1.787	5,375
Ethiopia	2000	154	-1.827	6,247	Rwanda	2010	64	-1.574	5,105
Ethiopia	2005	108	-1.502	5,680	Rwanda	2015	43	-1.457	4,798
Ethiopia	2011	82	-1.309	6,575	Senegal	1993	129	-1.172	2,774
Gabon	2000	92	-1.103	2,580	Senegal	2005	100	-0.793	5,683
Gabon	2012	74	-0.986	3,751	Senegal	2011	63	-1.197	6,157
Ghana	1993	88	-1.268	2,186	Senegal	2013	50	-0.935	3,202
Ghana	1998	108	-1.065	1,935	Senegal	2014	41	-1.040	3,277
Ghana	2003	89	-1.272	2,254	Tanzania	1992	125	-1.830	4,680
Ghana	2008	73	-0.829	1,807	Tanzania	1996	126	-1.799	4,023
Ghana	2014	52	-0.836	3,561	Tanzania	1999	147	-1.683	1,910
Guatemala	1995	69	-2.190	6,029	Tanzania	2005	88	-1.583	5,108
Guatemala	1999	65	-2.121	2,924	Tanzania	2010	81	-1.524	4,687
Guatemala	2015	33	-1.819	7,398	Tanzania	2015	58	-1.070	6,108
Guinea	1999	186	-1.013	3,148	Togo	1998	99	-1.228	3,832
Guinea	2005	135	-1.208	3,857	Togo	2014	77	-1.074	4,108
Guinea	2012	100	-0.743	4,130	Uganda	1995	154	-1.580	4,293
Haiti	1995	119	-1.293	2,092	Uganda	2001	131	-1.651	4,227
Haiti	2000	110	-1.055	3,988	Uganda	2006	106	-1.354	4,994
Haiti	2006	83	-1.151	3,640	Uganda	2011	80	-1.186	4,673
Haiti	2012	79	-0.887	4,436	Zambia	1992	180	-1.750	3,938
India	1993	92	-1.897	36,168	Zambia	1996	181	-1.790	4,358
India	1999	80	-1.868	32,699	Zambia	2002	172	-1.910	4,181
India	2006	59	-1.506	29,902	Zambia	2007	110	-1.495	3,939
Kenya	1993	109	-1.438	3,567	Zambia	2014	57	-1.461	7,821
Kenya	1998	92	-1.386	3,409	Zimbabwe	1994	88	-1.135	2,400
Kenya	2003	103	-1.238	3,402	Zimbabwe	1999	95	-1.024	2,208
Kenya	2009	76	-1.222	3,702	Zimbabwe	2006	98	-1.279	3,188
Kenya	2014	48	-1.072	12,339	Zimbabwe	2011	89	-1.296	3,560
Madagascar	1992	132	-2.113	3,219	Zimbabwe	2015	56	-0.970	3,601

Source: DHS data sets; calculation by authors.

B. Observed, simulated, and imputed z-scores by country

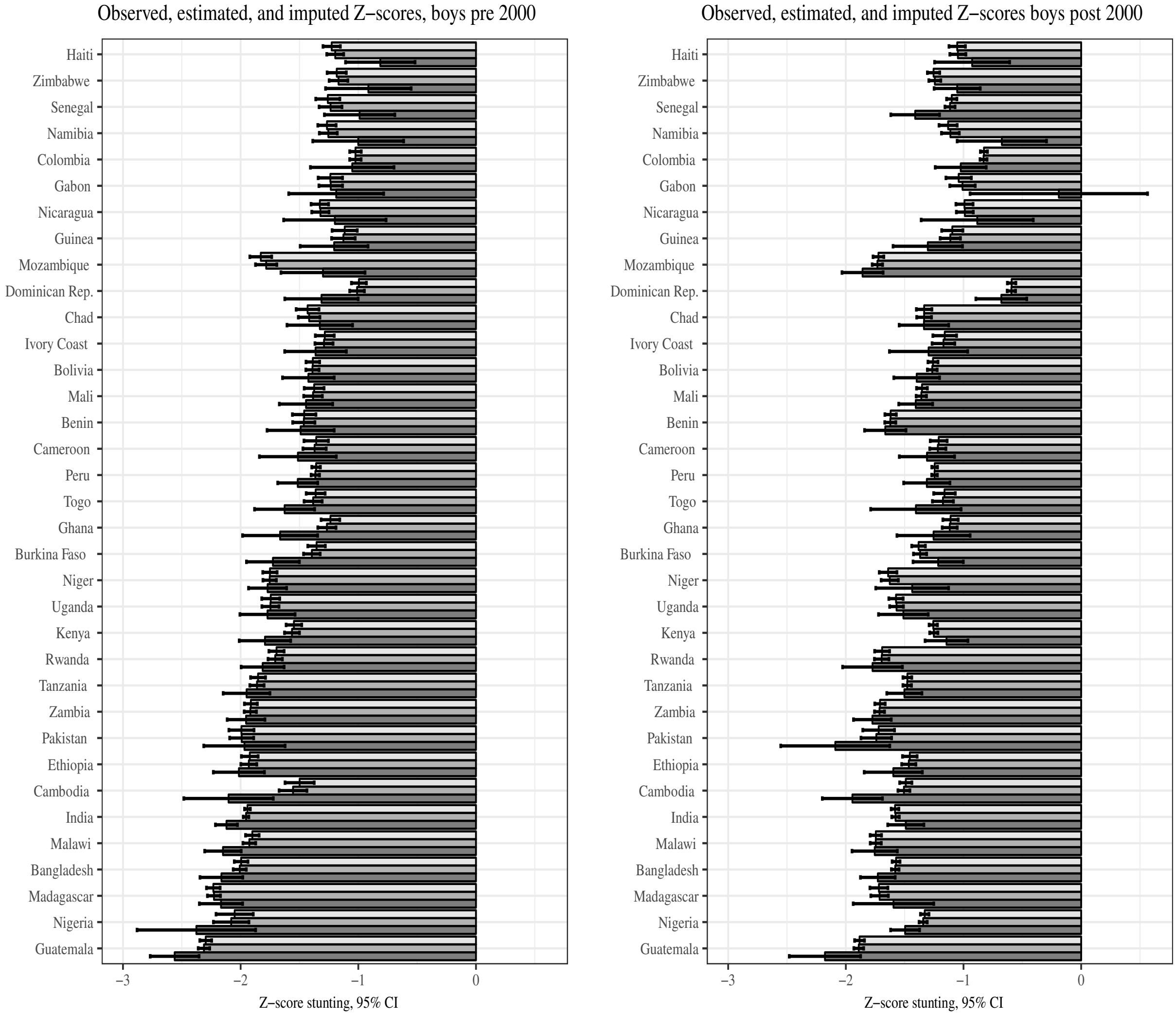


Figure B.1: Selective mortality and its impact on stunting for boys by country. The left-hand panel depicts the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores by country, together with 95% confidence intervals for the period before 2000. The right-hand panel depicts the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores by country, together with 95% confidence intervals for the period after 2000. **Source:** Own calculation with available data from DHS survey.

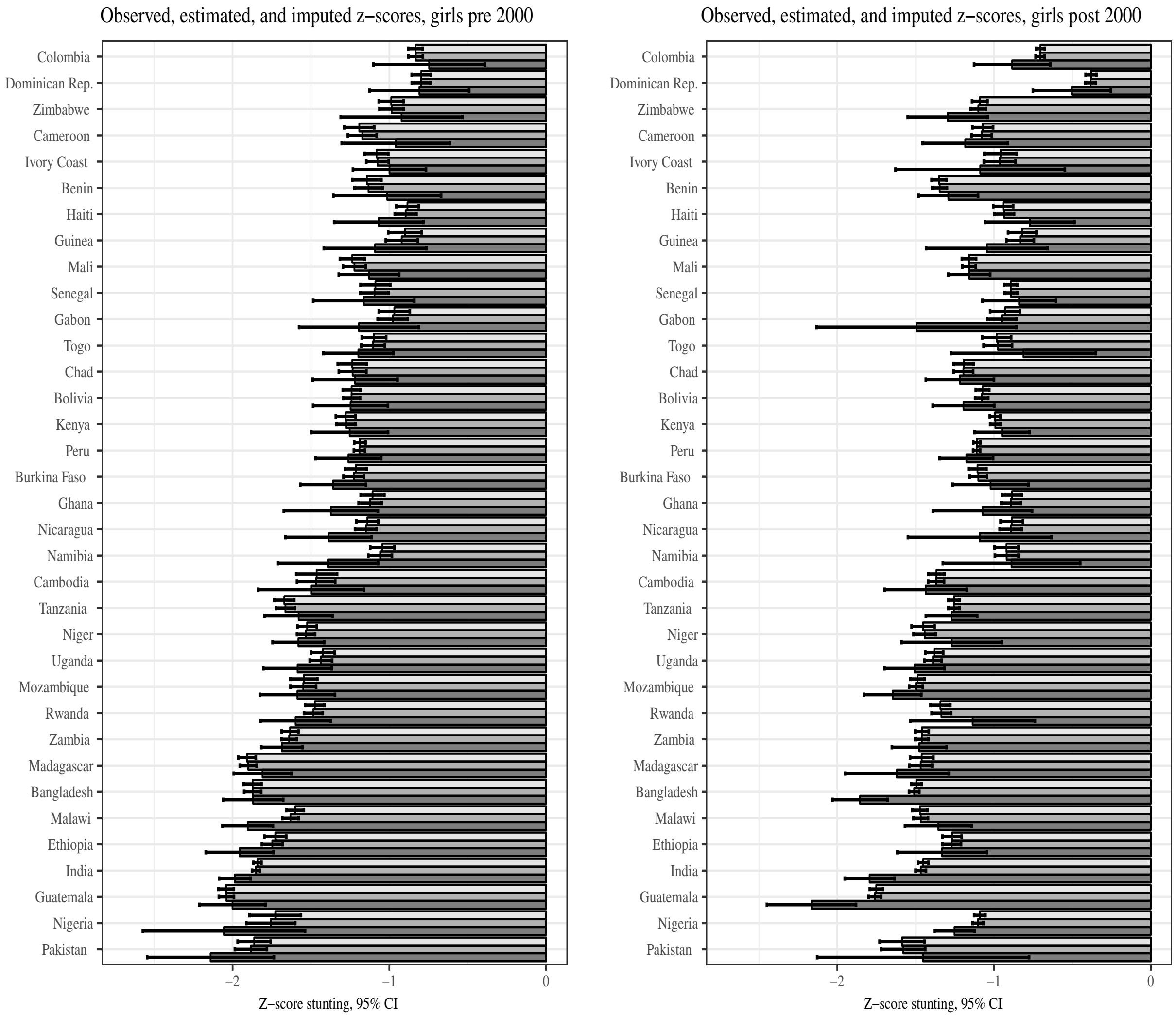


Figure B.2: Selective mortality and its impact on stunting for girls by country. The left-hand panel depicts the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores by country, together with 95% confidence intervals for the period before 2000. The right-hand panel depicts the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores by country, together with 95% confidence intervals for the period after 2000. **Source:** Own calculation with available data from DHS survey.

C. Robustness check

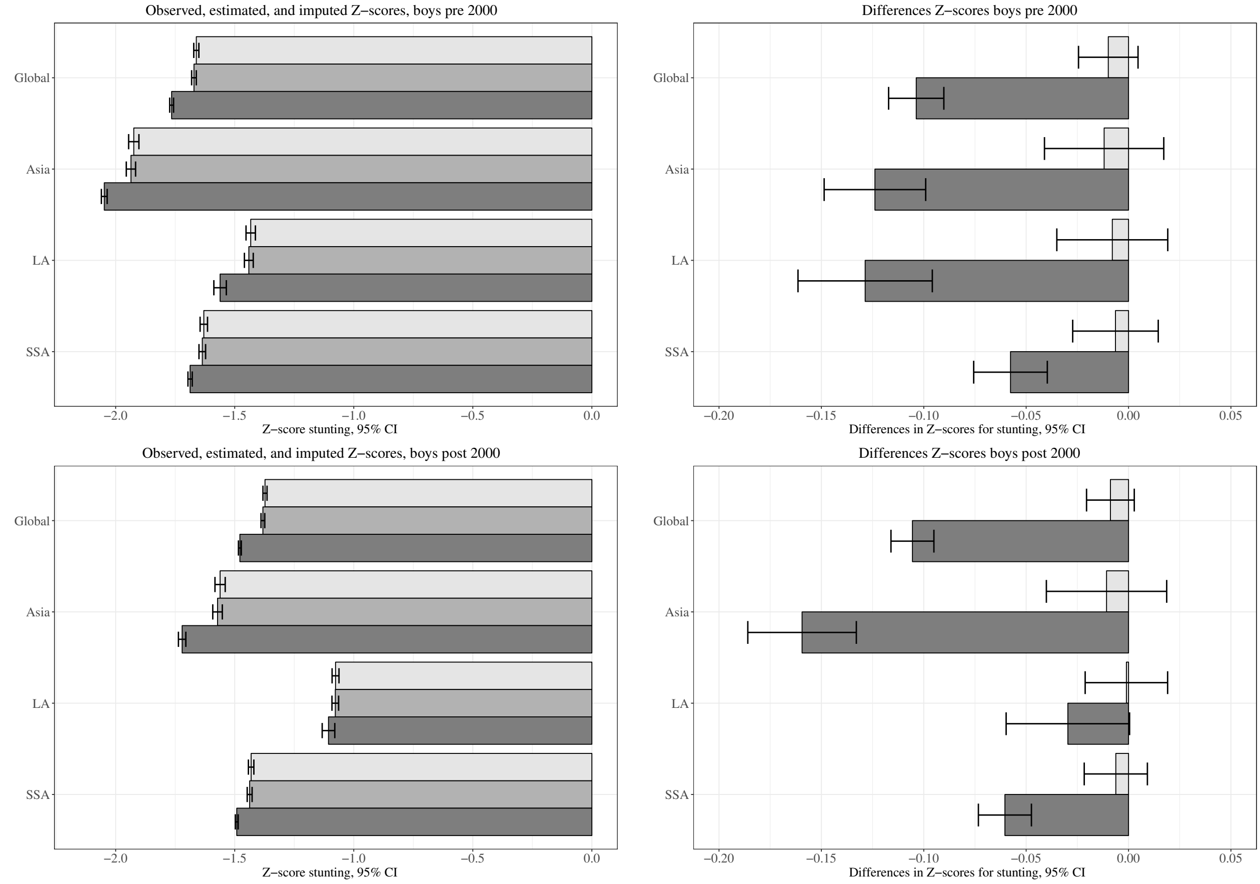


Figure C.1: Selective mortality and its impact on stunting for boys using out-of-sample prediction. The left-hand panels depict the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. The right-hand panels depict the difference between the observed and the simulated z-scores (**grey**) and between the observed and the imputed z-scores (**dark grey**) by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. **Source:** Own calculation with available data from DHS survey.

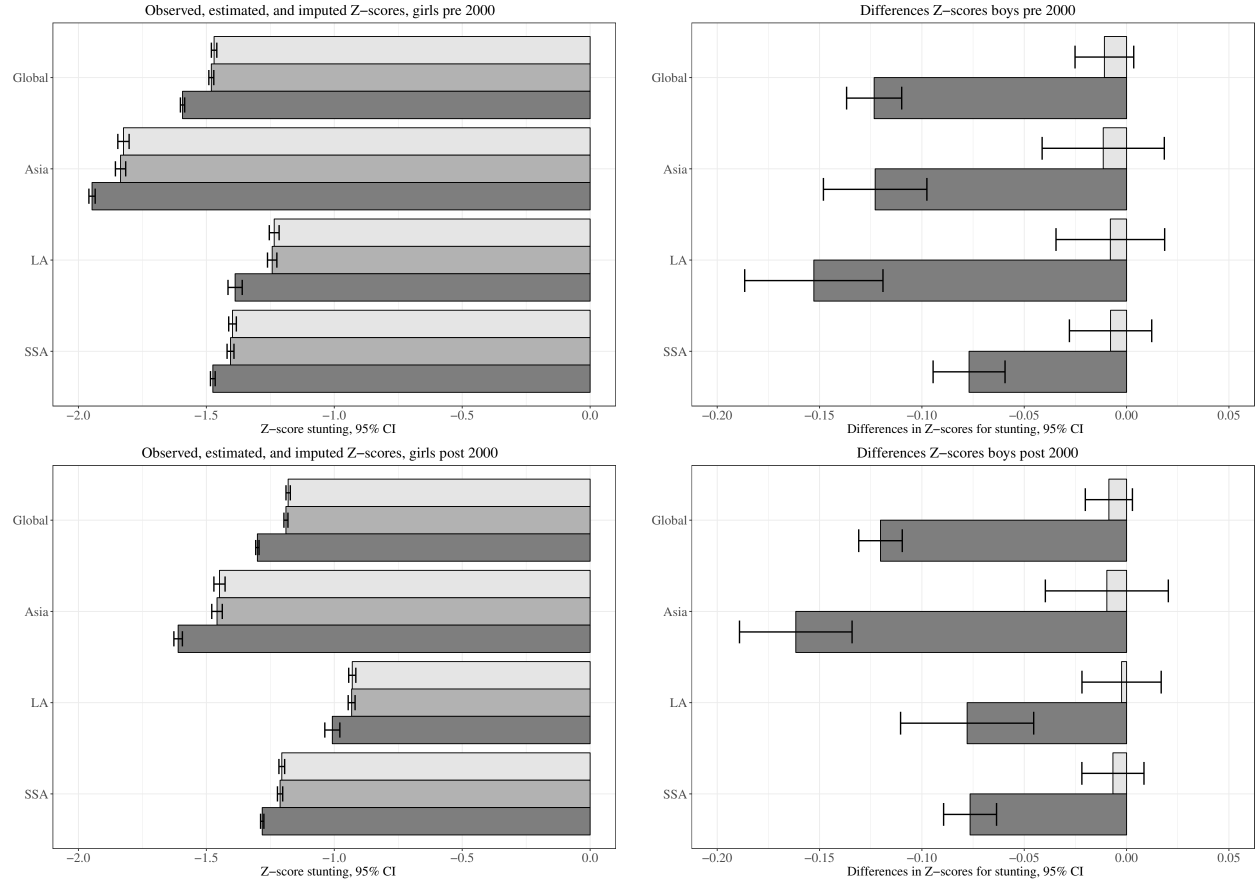


Figure C.2: Selective mortality and its impact on stunting for girls using out-of-sample prediction. The left-hand panels depict the observed (**light grey**), simulated (**grey**), and imputed (**dark grey**) z-scores by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. The right-hand panels depict the difference between the observed and the simulated z-scores (**grey**) and between the observed and the imputed z-scores (**dark grey**) by geopolitical region, together with 95% confidence intervals for the periods before and after 2000. **Source:** Own calculation with available data from DHS survey.

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

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

2017-27 **Kenneth Harttgen, Stefan Lang, Johannes Seiler:** Selective mortality and undernutrition in low- and middle-income countries

2017-26 **Jun Honda, Roman Inderst:** Nonlinear incentives and advisor bias

2017-25 **Thorsten Simon, Peter Fabsic, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Probabilistic forecasting of thunderstorms in the Eastern Alps

2017-24 **Florian Lindner:** Choking under pressure of top performers: Evidence from biathlon competitions

2017-23 **Manuel Gebetsberger, Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Estimation methods for non-homogeneous regression models: Minimum continuous ranked probability score vs. maximum likelihood

2017-22 **Sebastian J. Dietz, Philipp Kneringer, Georg J. Mayr, Achim Zeileis:** Forecasting low-visibility procedure states with tree-based statistical methods

2017-21 **Philipp Kneringer, Sebastian J. Dietz, Georg J. Mayr, Achim Zeileis:** Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season

2017-20 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior

2017-19 **Martin Geiger, Richard Hule:** The role of correlation in two-asset games: Some experimental evidence

2017-18 **Rudolf Kerschbamer, Daniel Neururer, Alexander Gruber:** Do the altruists lie less?

2017-17 **Meike Köhler, Nikolaus Umlauf, Sonja Greven:** Nonlinear association structures in flexible Bayesian additive joint models

2017-16 **Rudolf Kerschbamer, Daniel Muller:** Social preferences and political attitudes: An online experiment on a large heterogeneous sample

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 Puntscher, 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 Bannier, 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 Puntscher, 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

2017-27

Kenneth Harttgen, Stefan Lang, Johannes Seiler
Selective mortality and undernutrition in low- and middle-income countries

Abstract
Anthropometric indicators, in particular the height for a particular age, are found to be lowest in South Asia compared to other geopolitical regions. However, despite the close relationship between undernutrition and mortality rates, the highest mortality rates are concentrated in sub-Saharan Africa. By accounting for this survival bias, i.e. selective mortality, this discrepancy between the undernutrition rates between South Asia and sub-Saharan Africa should be expected to decrease. In addition, one can also ask whether undernutrition rates would differ without selective mortality. Using data stemming from six waves of Demographic and Health Surveys (DHS), we assess the impact of selective mortality on the anthropometric indicators for the children's height-for-age (stunting), weight-for-age (underweight), and weight-for-height (wasting) for a global sample of low and middle income countries between 1991 and 2015. Taking advantage of a matching approach, the effect of selective mortality for a cross-section of 35 developing countries is analysed. This approach allows values, originally stemming from non-deceased children, to be assigned for the otherwise non-observed anthropometric indicators of deceased children. These values are imputed under the counterfactual scenario that these deceased children would still be alive. The results are twofold: First, this approach reveals that the imputed values for deceased children for stunting, underweight, and wasting are significantly lower compared to the observed anthropometric indicators. Second, the difference between the observed anthropometric indicators, and the constructed overall anthropometric indicators are found to be only of negligible magnitude. Only assuming high mortality rates, or imputing the lower bound considered by the WHO as cutoffs for outliers, would alter the second finding.

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