

Harttgen, Kenneth; Lang, Stefan; Seiler, Johannes

Working Paper

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

Working Papers in Economics and Statistics, No. 2017-27

Provided in Cooperation with:

Institute of Public Finance, University of Innsbruck

Suggested Citation: Harttgen, Kenneth; Lang, Stefan; Seiler, Johannes (2017) : Selective mortality and undernutrition in low- and middle-income countries, Working Papers in Economics and Statistics, No. 2017-27, University of Innsbruck, Research Platform Empirical and Experimental Economics (eecon), Innsbruck

This Version is available at:

<https://hdl.handle.net/10419/184978>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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



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
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

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

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

Selective Mortality and Undernutrition in Low- and Middle-Income Countries

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

[†]ETH Zurich, Department of Humanities, Social and Political Sciences

[‡]University of Innsbruck, Department of Statistics

9th January 2018

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 for children below 3 years which stems 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 low- and middle-income 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 cut-offs for outliers, would alter the second finding.

Keywords Child mortality · Undernutrition · Selective Mortality · Asia · Latin America · SSA

JEL-Classification I15, I32, J13

Disclosure statement No potential conflict of interest was reported by the authors.

Acknowledgements The computational results presented have been achieved (in part) using the HPC infrastructure LEO of the University of Innsbruck. We thank Macro International Inc., for collecting the DHS, and providing access to the data. J.S. thanks Federico Rotolo for helpful comments on the implementation of the model into the **R**-package **parfm**, and acknowledges financial support from a scholarship by the University of Innsbruck. All remaining errors are our own.

1. Introduction

Progress in reducing mortality rates, as well as the reduction in the prevalence of undernutrition among children since the formulation of the Millennium Development Goals (MDG) has been observed throughout developing countries (You et al., 2015a,b; Black et al., 2013). Besides this, several authors (Pelletier et al., 1995; Pelletier & Frongillo, 2003; Black et al., 2008, 2013) found a strong and positive relationship between the nutritional situation of children and the mortality risk, and around 50% of all deceased children before the age of 5 can be attributed to undernutrition. However, highest mortality rates are observed in sub-Saharan Africa, whereas the prevalence of stunted children is found to be higher in Asian countries (Klasen, 2008; Paciorek et al., 2013; You et al., 2015b; Jayachandran & Pande, 2017). Assuming the relationship between undernutrition and mortality to be correct, after accounting for selective mortality, we would expect much higher undernutrition rates in sub-Saharan Africa, and the discrepancy between the undernutrition rates between South Asia and sub-Saharan Africa would decrease. Furthermore, 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, 2006; Grantham-McGregor et al., 2007), such that 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, see for instance, Borma et al. (1992); Pitt (1997); Dancer et al. (2008); Alderman et al. (2011). The first, and the second articles examine 17 and 13 surveys, respectively, from different developing countries. The latter two articles are case studies for Bangladesh and India, respectively. 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. Data from the Indian National Health Survey (NHS) for the years 1992/93, 1998/99 and 2005/06 is used to determine differences of non-deceased and deceased children's anthropometric indicators 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 anthropometrics of deceased children. In addition, to 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 (see for instance, Moradi, 2010; Gørgens et al., 2012).

By pooling 140 surveys from 35 countries between 1991 and 2015, and assessing the effect of selective mortality at a global level, and for the geopolitical regions individually, as well as individually for the included countries, we aim to contribute to the literature the following:

First, assess whether selective mortality exists, and determine the impact of selective mortality

on undernutrition rates.

Second, 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, as well as for the three geopolitical regions whether differences between the geopolitical regions exist. Additionally, we aim to identify whether the effect of selective mortality differs between countries by analysing selective mortality at a country level.

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 pre and post the year 2000. Therefore, the data consists of 140 individual surveys, between 1991 and 2015 from 35 Asian, Latin American, and sub-Saharan African countries. Using an approach similar as in [Alderman et al. \(2011\)](#), we assess the effect of selective mortality by pooling the three geopolitical regions with the largest share of low- and middle-income countries, as well as the corresponding regional sub samples. The analysis confirms the findings of earlier articles analysing selective mortality, and further expands the results to a global sample.

Fourth, we evaluate if a systematic underestimation of the effect of undernutrition is introduced by the joint analysis of child mortality and undernutrition. The source of this underestimation would be, as multilevel modelling becomes feasible, and more popular in development economics, several studies exist that analyse underlying causes of child mortality using this technique (see for instance, [Misselhorn & Harttgen, 2006](#); [Adedini et al., 2015](#); [Harttgen et al., 2015, 2017](#)), it is important to include the children's nutritional situation in the estimation. This is because child undernutrition is found to be one of the main underlying causes of child mortality (see for instance, [Pelletier & Frongillo, 2003](#); [Black et al., 2008](#)). However, anthropometric characteristics are observed by the Demographic and Health Surveys (DHS) which is one of the main data source, only for non-deceased children. Hence, considering the nutritional situation is not possible at an individual level. Therefore, a hierarchical model is particularly useful as it allows, as an approximation for undernutrition, an aggregation at the next higher level. However, this introduces a potential bias in the estimation, as it can be assumed that the true nutritional status of deceased children is overestimated by this approximation, which yields the effect of undernutrition on mortality to be underestimated. Evaluating this, we will assess whether a systematic underestimation of the effect of undernutrition is introduced by the joint analysis of undernutrition and child mortality. Therefore, we aim to identify the effect of selective mortality on the observed anthropometric indicators for a sample of low- and middle-income countries, to determine whether the effect yields significant changes in the anthropometric indicators. In addition, by analysing the relationship of selective mortality and undernutrition we try to get a better understanding of the underlying relationship, and aim to validate whether a influential bias is introduced when aggregating the anthropometric indicators at a regional or country level. By pooling 140 surveys

from 35 low- and middle-income countries, in addition, we aim to determine the effect of selective mortality on the anthropometric indicators for a broader sample over time.

The remainder of this paper is organized as follows: Section 2 will give an overview about the data set we used and will also give descriptive statistics for the anthropometric indicators. This section is followed by Section 3 which illustrates the potential impact selective mortality has on the anthropometric indicators. Section 4 will briefly illustrate the applied statistical method. Section 4 is followed by Section 5 and Section 6 where we present and discuss the results for the anthropometric indicators stemming from this approach. In the last section of our paper, Section 7, we will give concluding remarks to our findings for the anthropometric indicators.

2. Data and summary statistics

The sources of the data for our analysis are the countries' national DHSs. 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). Currently, the 7th wave is going on, and usually the spacing between waves is 5 years. The data to be collected is standardized across countries, and comprise data on topics such as child health, child mortality, child and maternal education, and maternal health to the 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. We included countries from the three mentioned regions for which at least one survey pre and post the year 2000 was available, to allow for an interpretation over time. Seven countries are located according to the classification by the DHS in Latin America & Caribbean (henceforth Latin America), four countries are located according to the classification made by the DHS 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 3 years of living, 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 pre and post the year 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% (boys), 33.7% (girls)), indicating that children born in sub-Saharan Africa are faced with a high mortality risk for the longest time. Whereas in the other geopolitical regions the share of children dying in the first month of living is between 45% and up to 60%, Asia (60.4% (boys), 53.5% (girls)), and Latin America (48.5% (boys), 46.7% (girls)) the exposure to the mortality risk declines more strongly. 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 previous to the year 2000, to -1.37 after the year 2000 for boys, and from -1.47 to -1.18 for girls, respectively. See Table II and Table 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](#)). Looking at the height-for-age z-score for the three analysed geopolitical regions individually, a large dispersion between the regions is 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. The same pattern emerges when comparing the stunting rates, defined by the WHO as not stunted, moderate stunted, and severely stunted, 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 respectively. When looking at the share of severely stunted children, the share of severely stunted children decreased over time. Looking at the geopolitical regions, the share of severely stunted children is found to be highest in Asia, however decreases from 28.3% to 18.9% for boys, and from 26.1% to 16.2% for girls, respectively. The lowest shares of severely stunted children are found in Latin America. However, also 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 additionally 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, respectively 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 (see for instance, [Pelletier et al., 1995](#)).

[Please insert TABLE II about here]

[Please insert TABLE III about here]

3. Selective mortality and potential impact on anthropometric indicators

Several studies found that undernutrition is one of the main underlying causes of child death, e. g. [Pelletier et al. \(1995\)](#) found in an analysis of 53 developing countries that around 56% of the observed deaths can be attributed to undernutrition. Several other studies found similar values for the role of undernutrition as the underlying cause of child mortality. For instance, [Bryce et al. \(2005\)](#) found that undernutrition is responsible for 53% of all deaths of children up to the age of 5 years. However, these studies either use country aggregates or use only small samples to estimate the effect of undernutrition on the risk of dying. This raises the question on whether the data is geographically representative to provide valid estimates. By using DHS data we overcome this critique and aim to provide better estimates for the observed relationship.

Other studies which emphasized the importance of the link between undernutrition and mortality are among others; [Rutstein \(2000\)](#); [Pelletier & Frongillo \(2003\)](#); [Black et al. \(2008\)](#). Most surveys, such as the DHS, only observe the anthropometric indicators at a specific point in time, and only for children alive at the day of the interview. Accordingly, if the aforementioned relationship exists at such a high level, the observed anthropometric indicators would not represent a valid cross section of the population, and due to selective mortality anthropometric indicators for the population would be overestimated.

Selective mortality. Since anthropometric indicators of the development status of children are only observed for living children, ignoring this issue and the effect of selective mortality could potentially yield to a bias in any further analysis. For instance, this is the case in a hierarchical model as the observed anthropometric indicator (Z_s) overestimates the ‘true’ anthropometric indicator for deceased and non-deceased children (Z_{all}). Reformulating this gives, as formulated by [Alderman et al. \(2011\)](#) the following relationship between Z_{all} , Z_s and Z_d the unobserved z-score of deceased children:

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

where P_d is the share of deceased children and Z_d the unobserved z-score of dead children. As mentioned previously a certain share of fatalities can be linked to undernutrition which gives the property of $Z_d < Z_s$. Neglecting this property could potentially lead to the overestimation of Z_{all} . However, since under-3 mortality rate per 1,000 live births for the corresponding period is about 100, the contribution of the observed z-score (Z_s) 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 which can be accounted to undernutrition the unobserved z-score of deceased children should be lower compared to the observed z-score. However, 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 Imputing values for the unobserved anthropometric indicators of children who died, to analyse the impact on the overall z-score (Z_{all}) under the counterfactual outcome, is done under a simple imputation scenario in a first step. Four different cases are analysed: First (I1), imputing a z-score for deceased children which results in statistical significant change ($\alpha = 0.05$) in the overall z-score. Second (I2), for the sample of deceased children the double of the observed average z-score is imputed. Third (I3), the lower bounds considered by the WHO as cut-offs to identify outliers for stunting are imputed¹. And fourth (I4), the upper bounds found by Pelletier et al. (1995) where imputed.

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%, obviously 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. Compare Table IV for detailed results under the four imputation scenarios.

[Please insert IV about here]

¹Contrary 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 $> +5.0$ ($+5.0$);

4. Methodology

The statistical analysis is conducted using **R** (R Core Team, 2016) and the **R**-package **parfm** (Munda et al., 2012), for estimating a parametric survival model with a individual frailty term. Matching is done using the **R**-package **StatMatch** (D’Orazio, 2016).

In the previous section a simple imputation which was not based on individual characteristics has been conducted. Depending on the scenario, the results for the overall z-score show a moderate decrease under three of the four different scenarios. The extreme case scenario, imputing the lower-bounds yield a more extreme decrease of the overall z-score. In the next step, since survival 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 ($h_i(t)$) for individual i , considering the baseline hazard $h_0(t) = \lambda\gamma t^{\gamma-1}$, and the vector of covariates \mathbf{x}_i is given by,

$$\begin{aligned} h_i(t) &= h_0(t) \exp(\mathbf{x}_i\beta) \\ &= \gamma t^{\gamma-1} \exp(\mathbf{x}_i\beta). \end{aligned} \tag{2}$$

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),

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

The regression parameters of the Weibull proportional hazard model can be estimated using maximum likelihood procedures. From these estimates the matching score will be derived, which allows values to be assigned for the unobserved anthropometric indicators of deceased children from 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 vital children. Matching follows the ideas of Rubin (1976, 1979); Rosenbaum & Rubin (1983), who introduced propensity score matching. However, 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, 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 assigned 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, and so to say the overlap in both groups is large enough. This procedure, usually first estimates the propensity score via a logit or probit regression. However, due to its ability to account for censoring, truncation a Weibull model is estimated. Additionally, the Weibull distribution can be used to model skewed data. As already pointed out earlier, deaths are not uniformly distributed in the first 36 months after birth. Clearly deaths are skewed-right with most death occurring between birth and the first month of living. 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).

5. Results

Descriptive statistics of the main explanatory variables are shown in Table V, for both the complete sample, and the regional sub-samples, for the periods pre and post the year 2000. Throughout all regions the share of female observations is always slightly below 50%. The average household size decreased in all regions when comparing pre to post 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 in general a better nutritional situation (see Table II and Table III). We find the results to be rather similar between the different three different anthropometric indicators, and only the results for stunting will be presented.

[Please insert TABLE V about here]

5.1. 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 for the unobserved anthropometric indicator of deceased children. The Weibull proportional hazard model includes characteristics of the children, characteristics of the mother and characteristics of the household. Unfortunately, the DHS do not report a subjective birth weight measure 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 of the observations. 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. However, the 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 looks in contrast to the results of Alderman et al. (2011), where they find an insignificant coefficient for gender, this discrepancy is accounted for. This is because their study is done for India, where the existing preference for male children counter balances this effect, and under-5 rates are lower for girls than for boys in countries where gender preferences are insignificant (Arnold, 1997). In addition to 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, and 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, 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]

5.2. Results imputing anthropometric indicators

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, pre and post the year 2000. Figure 2 illustrates the results for boys, while Figure 3 illustrates the results for girls. In more detail, the left-hand panel depicts the observed z-scores, the constructed overall z-scores, and imputed z-scores of deceased children. Whereas in the right-hand panel the differences between the observed z-scores and the constructed overall z-scores, and the differences between the observed z-scores and the imputed z-scores are illustrated, respectively. The observed z-scores and constructed overall z-scores are more or less identical, and the difference is never statistical significant. However this, is not surprising, 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 per 1,000 live birth below 100, 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 between boys and girls for the difference between observed z-scores and constructed overall z-scores is negligible. Nevertheless, the result of this comparison as already done with data for India by Alderman et al. (2011) should not be surprising, as the vast majority of the values used to construct the new overall z-scores stem from the observed z-scores.

Naturally, it is more interesting 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 the year 2000 the difference is around 0.05 for both boys and girls. Whereas to our surprise, the difference increases to around 0.075 for boys and around 0.1 for girls after the year 2000. In addition, to these major findings we also find that the difference varies across geopolitical regions. For South Asia the differences pre and post the year 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 is of particular interest, as 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 found to be higher in South Asia, whereas in sub-Saharan Africa higher mortality rates are observed (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. However, 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. To get a better understanding of the effect of selective mortality on undernutrition, in addition to the main analysis illustrated in the previous section, in addition the effect of selective mortality on undernutrition is analysed at a country level for the 35 countries included in the analysis. Figure 4 and Figure 5 depict the difference between the observed and simulated z-score, Figure B.1 and Figure B.2 of the Appendix depict the observed, constructed, and imputed z-scores individually for the 35 countries pre and post the year 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 results for the included countries individually, differences between observed z-scores, and constructed z-scores the same pattern emerges as for the global sample. 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.

Comparing the observed height-for-age z-scores with the imputed z-scores of deceased children no clear pattern emerges. Meaning that differences across geopolitical regions vary strongly and structure between geopolitical regions can be seen. Previous to the year 2000 around 20%-30% of the included countries show lower, and significant imputed z-scores, whereas for the vast share of countries show lower but insignificant imputed z-scores, and only a minor share of the countries show higher imputed z-scores in comparison to the observed z-scores. After the year 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 as for boys. This finding is consistent with the analysis for Bangladesh of [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 context of this, we find for countries with a clear and strong male preference like Bangladesh, or India (see 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 bigger for girls, in comparison to the differences observed by boys. Within sub-Saharan Africa this pattern was not observed which is not surprising, as a gender preference for boys exists only in a few sub-Saharan African countries and is only weak (Khera et al., 2014; Rossi & Rouanet, 2015).

[Please insert FIGURE 4 about here]

[Please insert FIGURE 5 about here]

Robustness check using out-of-sample prediction. As a robustness check, and as an alternative approach to assess the difference between observed and unobserved z-scores of deceased children, out-of-sample prediction for the z-scores can be used. By doing so, and to gain a maximum of information from the metric covariates, we append the usually assumed linear predictor η_i^{lin} for the response, with the more general additive predictor η_i^{add} , where the metric covariates will be approximated using penalized splines. The corresponding semi parametric 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):

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

As the nutritional status of the child (z-score) is measured continuously we assume the response to be normally distributed with variance σ^2 ($\mathcal{N}(\mu_i; \sigma^2)$), and will model the response as a function of socio-economic, maternal, and child specific characteristics of undernutrition, see Fenske et al. (2013); Lang et al. (2014) for reference studies focusing on modelling socio-economic, maternal, and child specific characteristics of undernutrition. $f(z_{i1})$ 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, and to account for a time trend the year the survey was conducted. In addition, we include country random effects. 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 semi parametric models was first introduced 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:

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

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

Results for this approach are comparable to the illustrated results of the matching approach, however the differences between the observed z-scores, and the imputed z-scores tend to be larger and can be summarized as follows:

- Imputed z-scores pooling the three geopolitical regions are between 6.0% and 10.0% lower compared to the observed z-scores (6.09% pre 2000 boys, 7.60% pre 2000 boys; 7.95% pre 2000 girls, 9.84% pre 2000 girls),
- the difference between the observed and imputed z-scores are similar when comparing between boys and girls,
- the difference between observed and imputed z-scores are lower in the period post the year 2000 (compared to the period before the year 2000)
- the difference between observed and imputed z-scores are in sub-Saharan Africa smaller (2.87% pre 2000 boys, 4.40% pre 2000 boys; 4.26% pre 2000 girls, 6.31% pre 2000 girls) then in South Asia (6.60% pre 2000 boys, 9.92% pre 2000 boys; 6.87% pre 2000 girls, 10.35% pre 2000 girls).

For an more elaborate overview of the results using out-of-sample prediction to predict the z-scores of deceased children see Appendix C.

6. Discussion

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 on average the observed z-score increased over time, the difference between the observed and unobserved z-scores decreased, showing to some extent

improvements towards achieving the Sustainable Development Goal (SDG) 2².

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. However, the prevalence of undernutrition is still found to be high (Paciorek et al., 2013), and in addition up to our knowledge one study exists (Sachdeva et al., 2016) which was conducted with hospitalized children between 6 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 over time, 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 evolved differently over time: India, despite a still high prevalence of undernutrition, experienced an improvement over time. In Ghana, despite on average a better nutritional status than India is observed, no improvements of the nutritional status was observed. Nicaragua showed - as the majority of 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 (see 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. This was also highlighted by the aforementioned study, however 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 the year 2000, where the difference between observed, and imputed stunting rates is actually higher for girls. This anomaly, for rates for girls in India, can be interpreted as a result of the gender preference in India, which is known to have a strong preference for boys Fuse (2010); Khera et al. (2014). Because of this, the household resources are probably not evenly distributed between boys and girls equally, potentially explaining the staggering difference between the differences between observed and imputed stunting rates between boys and girls. This effect of a differing resource allocation was also suggested by Jayachandran & Pande (2017). However, the more general pattern

²SDG 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.

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 of 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 un-observed anthropometric indicators of deceased children on average should have improved, too. This similar trend like for the observed anthropometric indicators can assumed.

Fourth, in the aforementioned study by Sachdeva et al. (2016), anthropometric indicators of hospitalized children between 6 and 59 months are observed for non-deceased, as well as deceased children. However, no distinction between boys and girls has been made. The data for this study comes from a hospital where mainly the poor urban population of East Delhi, and urban people of western Uttar Pradesh are treated. As the age of the included children differs from our analysis, no quantitative comparison can be made. However, 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]

7. Conclusion

This paper estimates the effect of selective mortality on anthropometric indicators, 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 the year 2000 the proportion of stunted and malnourished children was indeed higher in South Asia, compared to other geopolitical regions. However, 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 6 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. However, 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, and 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. However, due to the current declining mortality rates (in comparison over time) (see for instance, [You et al., 2015a](#)), and the improvement of the nutritional situation of children ([Paciorek et al., 2013](#)) with less 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 uses 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 status 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 reveal, as a starting point, that selective mortality will only have a significant impact on the overall anthropometric indicators in the case when 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, and 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

- (2010). *WHO Anthro for personal computers, version 3.2.2, 2011: Software for assessing growth and development of the world's children*. WHO, World Health Organization, Geneva.
- 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. Technical report, 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.
- D’Orazio, M. (2016). *StatMatch: Statistical Matching*. R package version 1.2.4.
- 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.
- Görgens, 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 implications for 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.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- 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.
- Rutstein, S. O. (2000). Factors associated with trends in infant and child mortality in developing countries during the 1990s. *Bulletin of the World Health Organization*, 78:1256–1270.
- 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.
- 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 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 (2006). *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development*. World Health Organization, Geneva, Switzerland.
- 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.

8. Figures

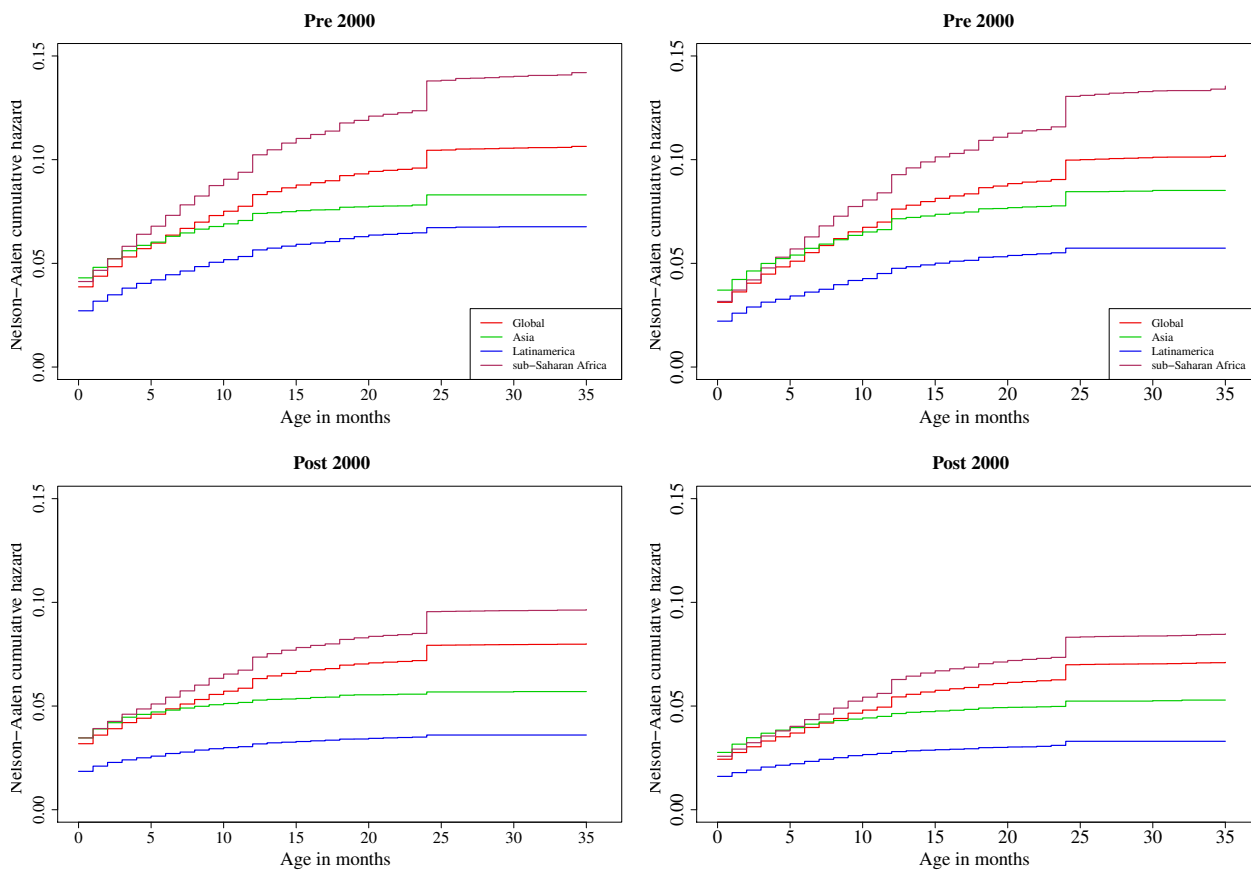


FIGURE 1: Nelson-Aalen cumulative hazard by gender, geopolitical region, and pre and post the year 2000. The left-hand panels depict the Nelson-Aalen cumulative hazard for boys, whereas the right-hand panels depict the Nelson-Aalen cumulative hazard for girls. **Source:** Own calculation with available data from DHS survey.

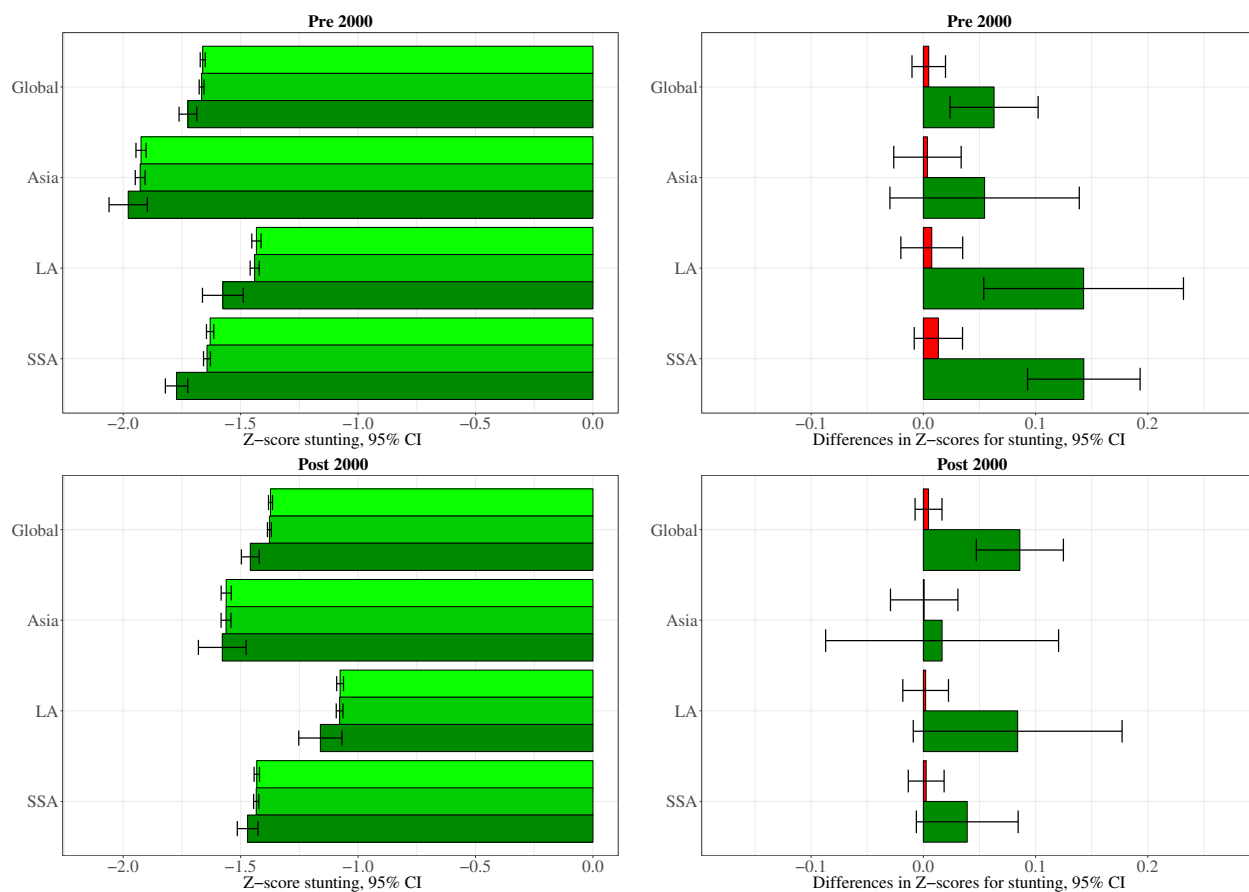


FIGURE 2: Selective mortality and its impact on stunting for boys. The left-hand panels depict the observed (light green), simulated (green), and imputed (dark green) z-scores by geopolitical region, together with 95% confidence intervals for the periods pre and post the year 2000. The right-hand panels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

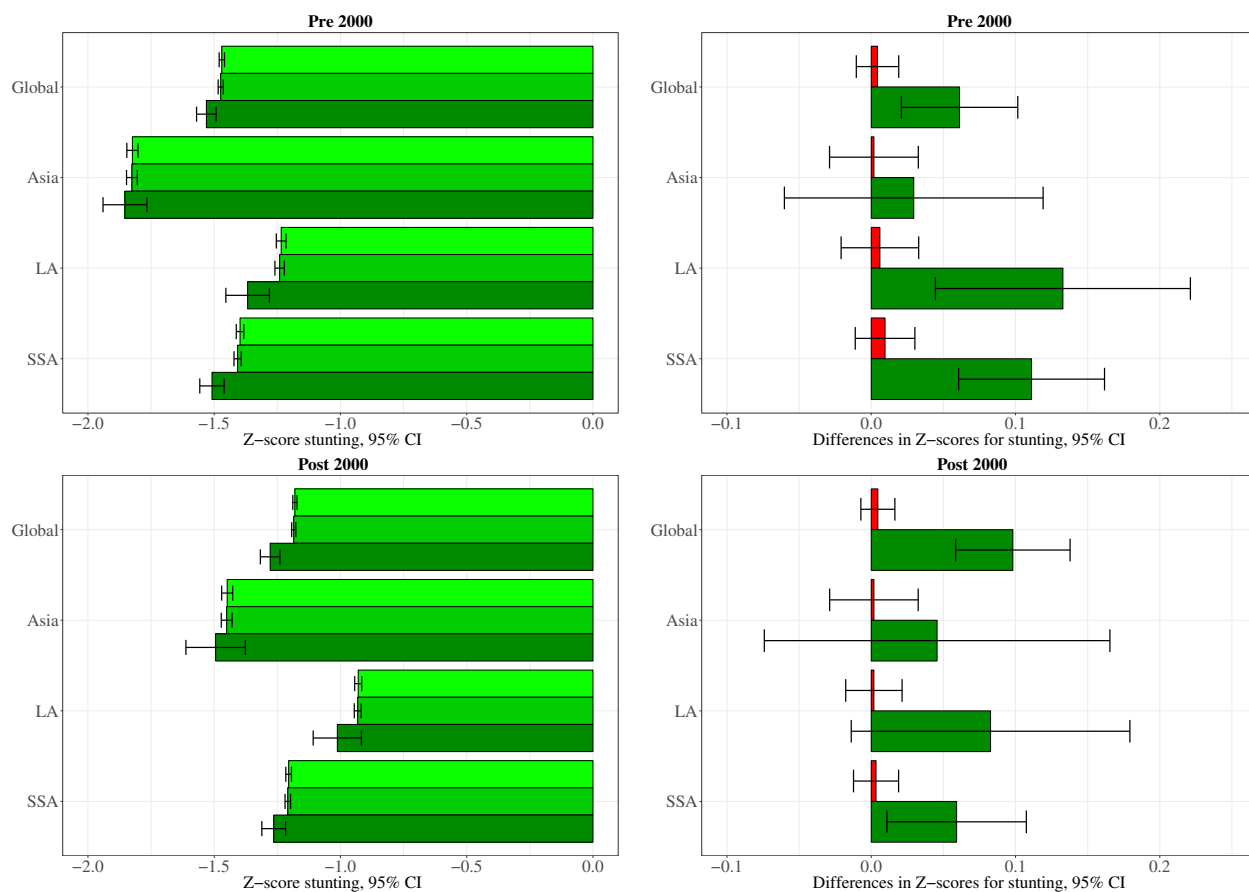


FIGURE 3: Selective mortality and its impact on stunting for girls. The left-hand panels depict the observed (light green), simulated (green), and imputed (dark green) z-scores by geopolitical region, together with 95% confidence intervals for the periods pre and post the year 2000. The right-hand panels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

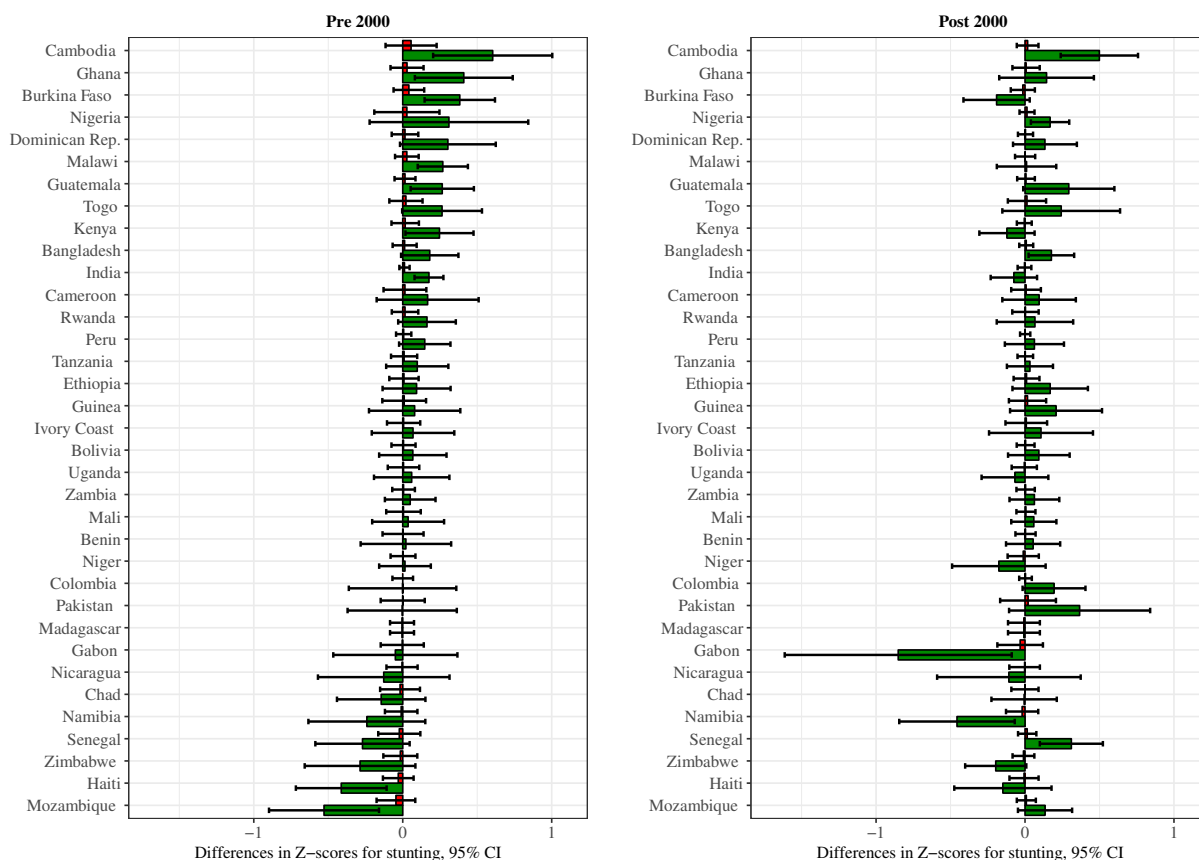


FIGURE 4: Selective mortality and its impact on stunting for boys by country. The left-hand panels depict the difference between the observed and the simulated z-scores (red) and between the observed and the imputed z-scores (dark green) by country, together with 95% confidence intervals for the period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

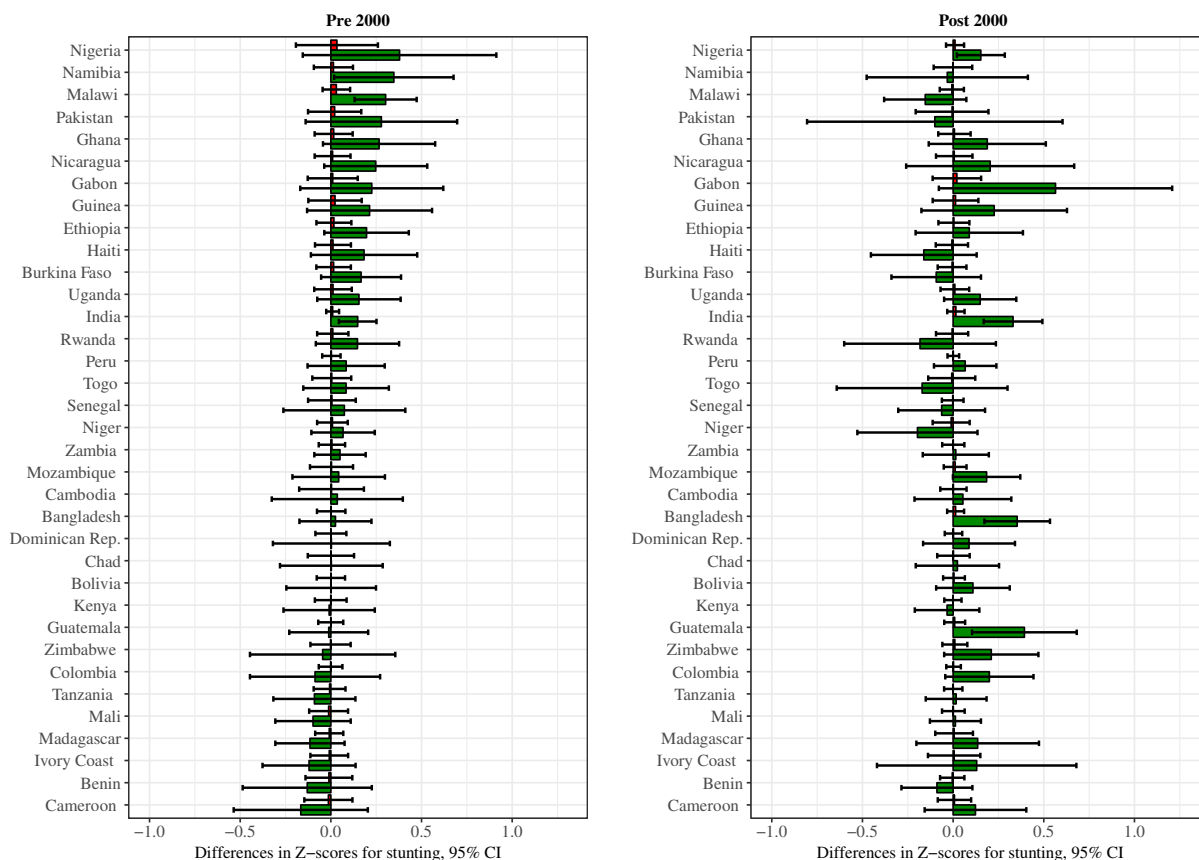


FIGURE 5: Selective mortality and its impact on stunting for girls by country. The left-hand panels depict the difference between the observed and the simulated z-scores (red) and between the observed and the imputed z-scores (dark green) by country, together with 95% confidence intervals for the period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

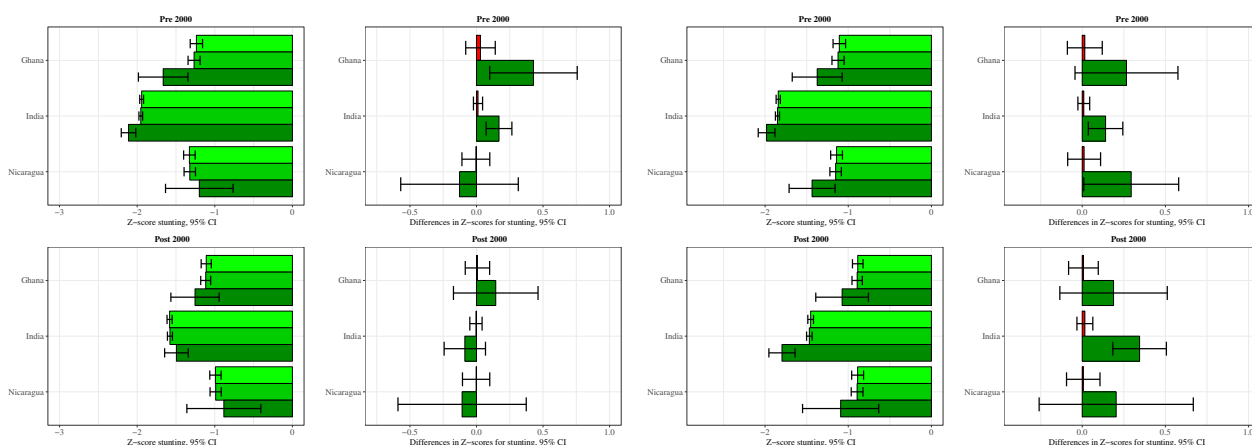


FIGURE 6: Selective mortality and its impact on stunting for selected countries. The left-hand panels depict the observed (light green), simulated (green), and imputed (dark green) z-scores for boys, together with 95% confidence intervals for the periods pre and post the year 2000. The right-hand panels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

9. Tables

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

Period	Age	Boys				Girls			
		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
Pre 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		Boys				Girls			
		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	Boys				Girls			
		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
	severly	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
	severly	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 ≤ -2 and < -3 ; and severely if the z-score is ≤ -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	Boys				Girls				
		Z_d	Z_s	Z_{all}	Δ	Z_d	Z_s	Z_{all}	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%	
	Eighth 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%	
	Eighth 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 3 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
Fifth 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(θ)	3.9892***	0.064	3.5821***	0.096	3.7652***	0.176	3.9446***	0.084
ln(ρ)	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
Fifth 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(θ)	4.2938***	0.064	-0.2383	0.302	2.5519***	0.229	4.258***	0.063
ln(ρ)	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

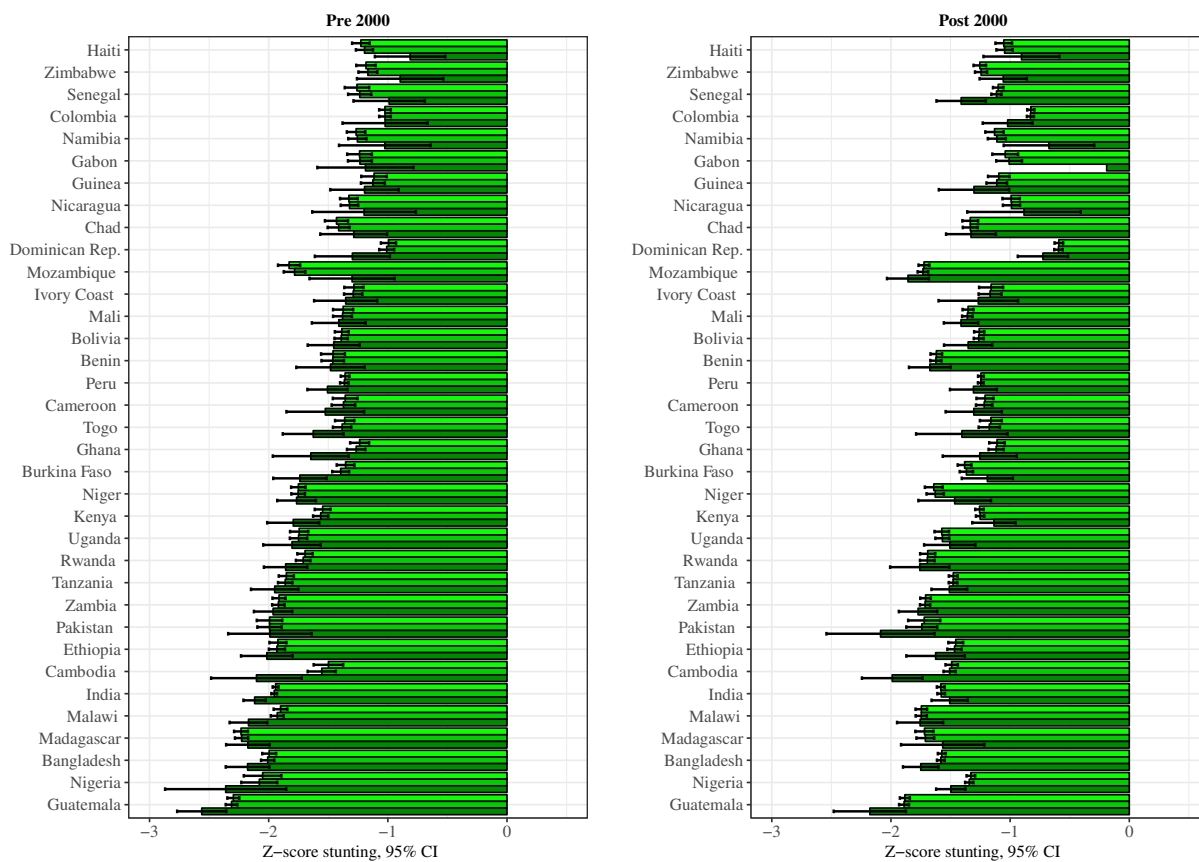


FIGURE B.1: Selective mortality and its impact on stunting for boys by country. The left-hand panel depicts the observed (light green), simulated (green), and imputed (dark green) z-scores by country, together with 95% confidence intervals for the period before the year 2000. The right-hand panel depicts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

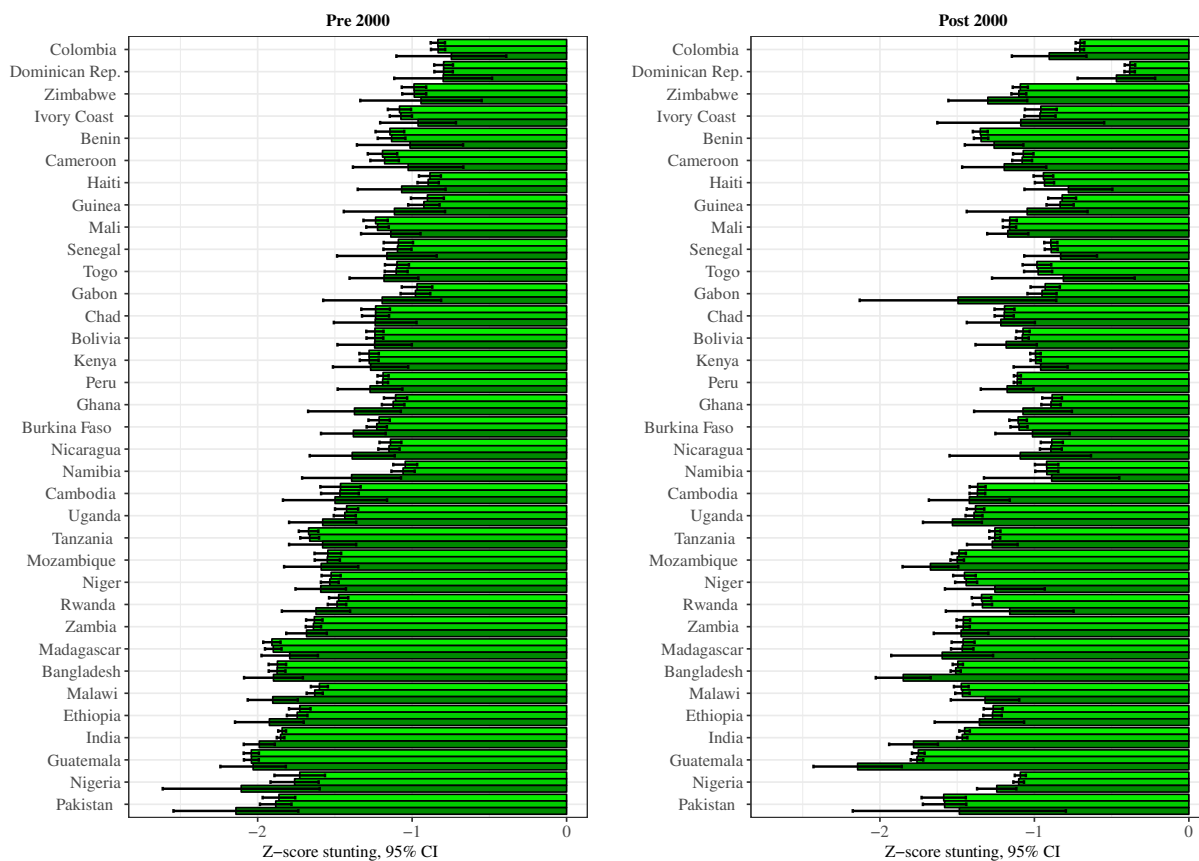


FIGURE B.2: Selective mortality and its impact on stunting for girls by country. The left-hand panel depicts the observed (light green), simulated (green), and imputed (dark green) z-scores by country, together with 95% confidence intervals for the period before the year 2000. The right-hand panel depicts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

C. Robustness check

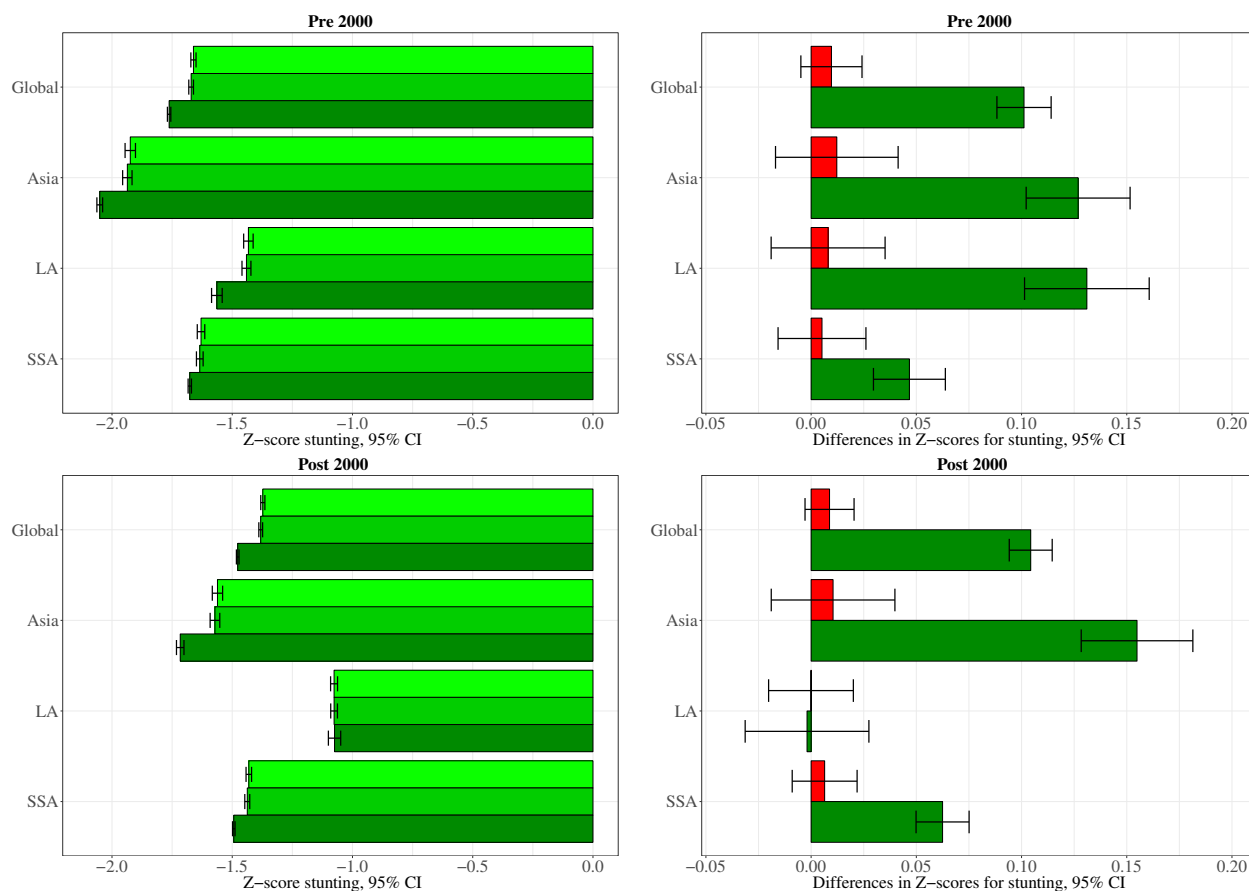


FIGURE C.1: Selective mortality and its impact on stunting for boys using out-of-sample prediction. The left-hand panels depict the observed (light green), simulated (green), and imputed (dark green) z-scores by geopolitical region, together with 95% confidence intervals for the periods pre and post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

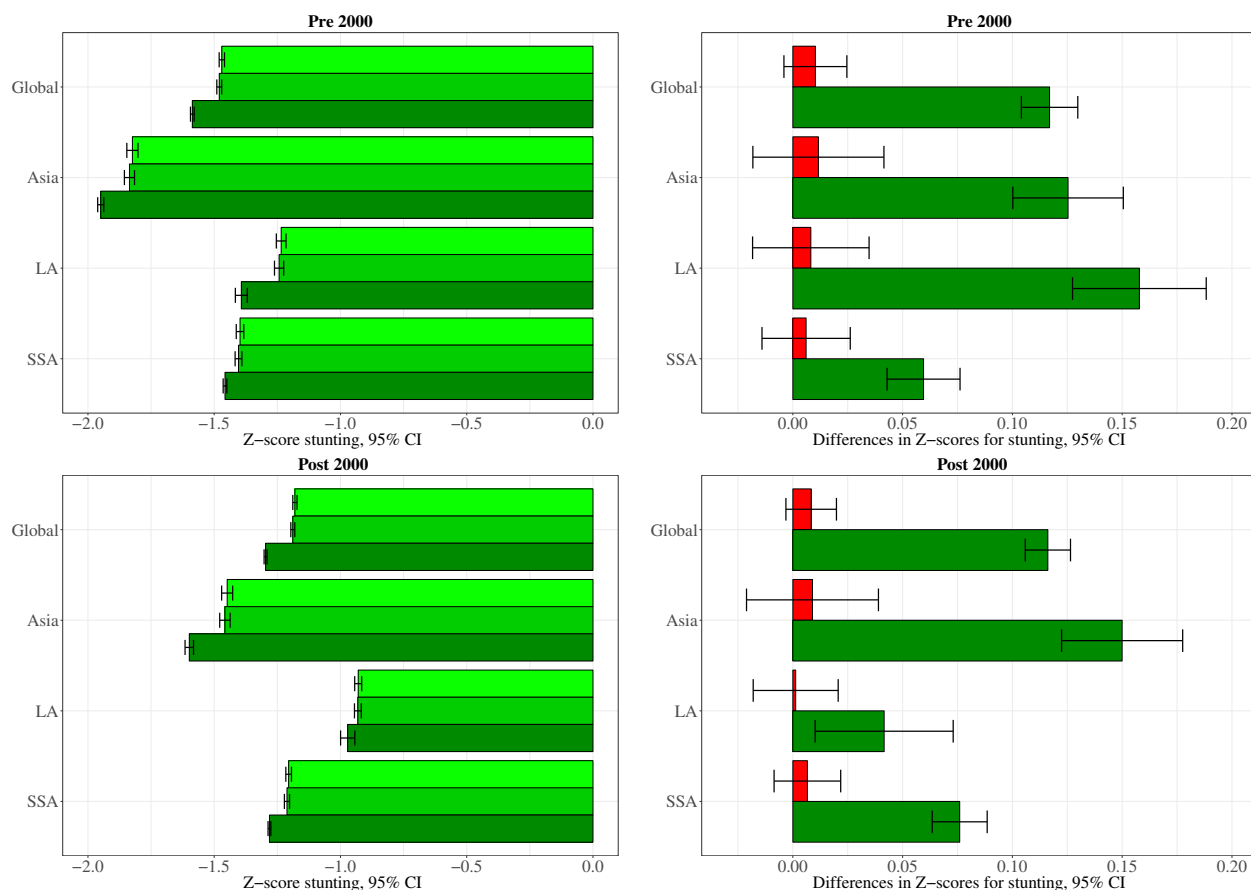


FIGURE C.2: Selective mortality and its impact on stunting for girls using out-of-sample prediction. The left-hand panels depict the observed (light green), simulated (green), and imputed (dark green) z-scores by geopolitical region, together with 95% confidence intervals for the periods pre and post the year 2000. The right-hand panels depict the difference between the observed and the simulated z-scores (red) and between the observed and the imputed z-scores (dark green) by geopolitical region, together with 95% confidence intervals for the periods pre and post the year 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 Puntcher, 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 Puntischer, 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)