Low Malnutrition but High Mortality: Explaining the Paradox of the Lake Victoria Region*

Johannes Gräb† and Jan Priebe‡

February 13, 2009

Exploiting DHS data from 235 regions in 29 Sub-Saharan Africa countries, we find that the combination of low levels of malnutrition together with dramatically high rates of mortality, encountered in Kenya's Lake Victoria territory, is unique for Sub-Saharan Africa. This paper explores the causes of this paradox for the Kenyan context. Our identification strategy consists of two parts. First of all, we apply multilevel regression models to control simultaneously for family and community clustering of the observed malnutrition and mortality outcomes. Secondly, to address unobserved but correlated factors, we exploit information from GIS and malaria databases to construct variables that capture additional components of children's geographic, political and cultural environment. Our analysis reveals that beneficial agricultural conditions and feeding practices lead to the observed sound anthropometric outcomes around Lake Victoria. In contrast, high mortality rates rest upon an adverse disease environment (malaria prevalence, water pollution, HIV rates) and a policy neglect (underprovision of health care services). Nonetheless, a significant effect of the local ethnic group, the Luo, on mortality remains.

Key words: Child mortality, undernutrition, poverty, multilevel modeling, Sub-Saharan Africa.

***JEL* codes:** I10, I30, O12, R12.

*We thank the Development Economics Research Group and the Center for Statistics in Göttingen for valuable comments. The paper benefited from discussions with Michael Grimm, Stephan Klasen, Stefan Sperlich and Walter Zucchini.

† University of Göttingen, Center for Statistics and Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany. Email: jgraeb@uni-goettingen.de.

‡ University of Göttingen, Center for Statistics and Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany. Email: jpriebe@uni-goettingen.de (corresponding author).

1 Introduction

There are a small number of regions in Sub-Saharan Africa that show very high levels of mortality given the anthropometric status of their population. Among these regions Nyanza, the principal Kenyan Lake Victoria province, is an exception of its own. In no other region in Sub-Saharan Africa (SSA) is the pattern of low levels of malnutrition together with dramatically high rates of mortality as pronounced as in Nyanza. Furthermore, the unique position of Nyanza is not only puzzling for Sub-Saharan Africa, but as well in the Kenya specific context. While Nyanza ranges on the upper limit of the mortality scale, most other Kenyan provinces depict comparatively low mortality rates given their levels of malnutrition.

In this paper we investigate the role of cultural, geographic, and political factors on the relationship of anthropometric outcomes of children and under-5 mortality rates in Kenya with an explicit focus on the unique situation of Nyanza and the territory around Lake Victoria. In order to disentangle the underlying mechanism that lead to the observed outcomes we analyze the factors driving mortality, stunting, and wasting jointly. Since parameter estimation can seriously suffer from endogeneity problems we adopted 3 strategies to mitigate this problem. First of all, we estimate reduced form regressions and therefore exclude any explanatory variable that we would expect to cause problems of simultaneous causality. Secondly, we augmented our DHS data by generating appropriate variables on malaria, health provision and Lake Victoria in order to mitigate problems arising from potential omitted variable bias. Thirdly, we use mixed model representations to further address unobserved heterogeneity issues on the family and community level.

Our findings point to a unique interaction of cultural, geographic and political factors in the Lake Victoria region which are responsible for causing the described paradox. Particularly, high mortality rates are found to rest upon the disease environment in the territory in combination with unfavorable cultural habits of the local ethnic group with respect to sexual, and pre- and post natal behavior. Political discrimination against this group resulting in reduced access to health infrastructure further exacerbates the mortality situation in the region. Nonetheless, even after controlling for other factors a significant ethnic specific influence on mortality remains although the effect is much smaller than found in previous studies. On the other hand, the area around Lake Victoria displays extraordinary positive conditions - fertile soils, a high level of food security and high protein availability (fish) - that contribute to children's advantageous nutritional outcomes.

In this regard the existing study adds an important new example and further insights to the few existing mortality-malnutrition paradoxes. For instance, the famous and widely investigated South Asia vs. Sub-Saharan Africa enigma (Ramalingaswami et al., 1996; Svedberg, 2000; Harttgen and Misselhorn, 2006; Klasen, 2008) refers to the observation that anthropometric outcomes of children are on average much better in SSA than in South Asia, while on the other hand child mortality rates are significantly higher in SSA compared to

South Asia.[1] Our study contributes to this literature for a variety of reasons. First of all, we are the first to explicitly state the Lake Victoria paradox and to analyze it comprehensively. Secondly, the study illustrates that not only in the context of potential large genetic differences (Klasen, 2008) such a paradox can arise, but that an interaction of cultural, geographic and political factors can reverse the positive relationship between a good nutritional status and the survival chances of children. Thirdly, in contrast to previous empirical studies on the South Asia vs. SSA enigma we explicitly control for factors related to the disease environment, e.g. HIV and malaria, and cultural factors and therefore are in a better position to obtain unbiased coefficients of our estimates. Fourthly, we use recent advances in multilevel modeling techniques that allow for the estimation of 3-level models which enables us to separate effects working at the individual, household, and community level. Moreover, this study is to our knowledge the most complete and accurate one analyzing the current determinants of under-5 mortality and anthropometric outcomes in Kenya.

Examining the nature of the Nyanza anomalies is further interesting and important since this example seems to question main findings in a variety of academic disciplines, most notably epidemiology, health economics, labor economics and economic history. In epidemiology, malnutrition, in particular wasting, low weight-for-height, is considered to be the main driver of mortality in developing countries (Villamor et al., 2005; Fawzi et al., 1997) and historical Europe (Fogel, 1994). Pelletier et al. (1995) claim that wasting, which is positively related to mortality due to diarrhea, fever and breathlessness, is the underlying cause of more than 50% of all child deaths in the world. Moreover, Caulfield et al. (2004) find that a sound nutritional status of children lowers the likelihood to die from malaria while Villamor et al. (2005) show that among HIV infected children the risk of dying in early ages was significantly higher for children being wasted. Thus, children who suffer from malnutrition have a significantly higher risk exposure to mortality due to a lower resistance to illnesses. This individual relationship is usually assumed to hold even on a higher aggregate level. Thus, areas with high prevalence of malnutrition rates are expected to have higher mortality rates. Regarding the Kenyan context we find the exact opposite pattern. While rates of wasted children gradually decrease with proximity to Lake Victoria, mortality rates steadily increase reaching its peak in Nyanza.

In health and labor economics the interest of studying the growth process of children has more recently been motivated by the findings that taller populations are economically better-off, more productive, and live longer (Bozzoli et al., 2009; Deaton, 2008). This result can partly be attributed to the prevailing disease environment. If on the one hand the child growth process and therefore adult heights and on the other hand life expectancy are positively correlated with the lack of certain diseases then for instance a simple Beckerian type of quantity-quality trade-off models can explain higher incomes of taller popula-

[1] In another mortality-malnutrition paradox Williamson (1990) finds that during the British industrial revolution the population in urban areas suffered from much higher mortality rates than the rural population despite their much better anthropometric status.

tions through the human capital formation process. The relationship between adult height and economic well-being might therefore only be valid if adult height is at the same time a good proxy for the mortality environment of a certain area or a country. In the absence of this later condition the relationship might be seriously flawed and this is what we partly observe in the Kenyan context.

Furthermore, there exist several studies in the field of economic history that make inferences about economic conditions during a specific time period based on mean height measures of population or population subgroups. Since height is expected to be an increasing but concave function of income, average height will be negatively correlated with initial income inequality (Steckel, 1995). Hence, income inequality has an effect on the dispersion of heights, so that inequality in height might function as an indicator of income inequality in the absence of data on the latter one, while mean height might serve as an indicator of mean income (Deaton, 2008). Obviously this inference does only hold if there are no other third factors that alter the relationship in an important way. Although this point has been recognized in the relevant literature (Deaton, 2008), it is often neglected or downplayed due to a lack of data that can function as control variables. Again the case of Nyanza and Lake Victoria illustrates that such an inference can simply go wrong.

The paper is structured as follows: Section 2 amplifies and describes the extent of the paradox of Nyanza and Lake Victoria in the Kenyan and Sub-Saharan African setting with respect to stunting, wasting and under-5 mortality. Section 3 discusses the theoretical model and outlines the identification strategy. Section 4 provides a detailed literature review on the most relevant cultural, geographical, and political particularities of the Kenyan context in order to explain the construction and interpretation of additional variables not included in DHS surveys. Section 5 provides descriptive statistics on the data sets and variables used in this study and comprises the multivariate analysis. Section 6 summarizes and concludes.

2 The Paradox

Evidence from the health and demographic literature suggests that there exists a clear positive relationship between a good nutritional status of a child and its chances of survival. From this observation it is usually inferred that this individual relationship holds as well at a higher aggregate level. Hence, regions or countries that perform well in terms of anthropometric indicators should exhibit lower mortality rates and vice versa.

However, the validity of this inference seems to be seriously challenged considering the spatial distribution of malnutrition and mortality in Kenya. Comparing the anthropometric and mortality outcomes of children in the Nyanza province to the other Kenyan provinces two things are salient. First of all, children in Nyanza score very well regarding anthropometric indicators, while infant and child mortality rates are extremely high in the region. Secondly, the mere extent of the within country variation in mortality rates astonishes. The extraordinary situation of Nyanza or more precisely the Lake Victoria region in the Kenyan context has been noted already for a long time. An investigation of 16 villages in 1922 in the area north of Lake Victoria showed an infant mortality rate between 335 and 514 per 1,000 live births with a tendency of declining infant mortality with distance to the lake (Colony and Protectorate of Kenya, 1923). Meanwhile historical data on adult mean height for the same time period suggests that the ethnic group living on the shores of Lake Victoria were the tallest in all over Kenya (Moradi, 2009).

Interestingly, although some authors occasionally have mentioned either the high mortality rates or the favorable anthropometric outcomes on the shores of Lake Victoria, to our knowledge no study exists that combines these two findings. Therefore, the puzzling situation on the shores of Lake Victoria has not yet been stated as such in the literature. Moreover, it has not been clear whether the observed mortality pattern combined with favorable anthropometric outcomes on the shores of Lake Victoria is even unusual for a larger geographical context. Hence, in order to assess the peculiarity of the Kenyan Lake Victoria region in the Sub-Saharan Africa setting we compiled a data set gathering information on anthropometric (prevalence of stunting and wasting) and mortality indicators (under-5 mortality rate) on regional level for all SSA countries where appropriate Demographic and Health Survey (DHS) data was available. If more than one DHS round was available for a country, we chose to only take the latest round into consideration. Furthermore, since the child growth reference standard which is used to calculate z-scores had changed rather recently, anthropometric statistics in our sample obtained from the official DHS reports would be based on two different standards. For the sake of comparison we recalculated stunting and wasting prevalence rates using the new WHO child growth reference standard for all regions (WHO, 2006). The final data set comprises 235 regions in 29 SSA countries.

Table 1 depicts under-5 mortality, stunting, and wasting rates for all regions with an under-5 mortality rate of 200 per 1,000 live births or more. Of these 36 regions only Tambacounda province in Senegal shows better stunting rates than the Nyanza province while with respect to wasting only Tete and Niassa provinces in Mozambique show lower prevalence rates. Moreover, it is interesting to note that all regions in the table that achieve prevalence rates in one of the anthropometric indicators similar to those in Nyanza, score much worse than Nyanza in the other anthropometric indicator. Therefore, Nyanza seems to be the only region in SSA that scores extremely well in stunting and wasting given the level of mortality.

[insert Table 1]

Even more striking is the result when focusing on the area in close proximity to Lake Victoria. Using the 2003 DHS round for Kenya jointly with the provided GIS data we calculated under-5 mortality rates and the two anthropometric indicators for all children born within a distance of 20km from Lake Victoria. In this area under-5 mortality rates strongly increase by approximately 50% to 306 per 1,000 live births compared to the Nyanza average while stunting rates even fall to 26.6% with wasting rates increasing slightly to 3.9%. Furthermore, the extreme position of the Lake Victoria region and Nyanza can be nicely illustrated as in Figure 1 which depicts the bivariate relationship between stunting and under-5 mortality for all 235 regions and the Lake region as defined above. Figure 1 underscores the unusual high mortality level of Nyanza given its stunting rates and much more important the unique position of the Lake area in the SSA context with an overwhelmingly high under-5 mortality rate given the level of stunting. A similar conclusion can be derived from Figure 2 which presents the bivariate relationship between wasting and under-5 mortality.

[insert Figures 1 & 2]

Besides highlighting the extraordinary situation of the Lake Victoria region, all 3 figures show further puzzling results. In particular, the within country distribution of under-5 mortality rates is remarkable. While Nyanza is situated far on the upper bound of the under-5 mortality rates given its level of stunting or wasting, several Kenyan provinces find themselves on the opposite side, showing very low mortality rates given its level in the respective anthropometric indicators. This high divergence of mortality levels within one country is highly unusual even for the SSA context. Column 9 in Table 1 presents the coefficient of variation (CV) based on the separate calculation for each country. Out of the whole sample Kenya shows the highest values in the CV among all 29 countries indicating the highest level of dispersion given its level of mortality.

Interestingly, the Lake Victoria provinces of Uganda and Tanzania do not exhibit such an unusual pattern as Nyanza. While anthropometric outcomes for children are slightly worse in these provinces compared to Nyanza, under-5 mortality rates are substantially lower. Since geographical and epidemiological conditions seem to be similar among all provinces bordering Lake Victoria, one possible explanation for this finding points to the role of the government and racial cleavages affecting migration decisions and the provision of health care. Compared to Uganda and Tanzania internal migration on provincial level in Kenya is limited by the prevalence of strong ethnic reservations. Furthermore, as discussed under section 4.3 discriminatory practices against the local ethnic group seems to have led to an underprovision of health care services in Nyanza compared to most other Kenyan provinces. In contrast, in Uganda the capital Kampala is situated in close proximity to Lake Victoria and therefore we would not expect an underprovision of health care services on the Ugandan side of Lake Victoria. Regarding Tanzania ethnic based inequalities in the provision of health care services are rather low due to the pursued nation building policies and hence health care provision seems to be much more need oriented (Miguel, 2004).

The simultaneous appearance of very low levels of malnutrition together with tremendously high rates of mortality in Nyanza and in particular the Lake Victoria region is unique in the SSA context. Moreover, the appearance of this phenomenon in a national context of relatively low mortality rates is further puzzling and led us to call it 'The Paradox'. Trying to explain this paradox is the objective of this paper and will demand a detailed review of the particularities of the Kenyan context.

3 Theoretical Framework

In order to analyze the described paradox we estimate reduced forms of child health and mortality production functions. The choice of our theoretical model relies on earlier work in this field done by Akin et al. (1992); Rosenzweig and Wolpin (1988) and in particular Behrman and Deolalikar (1988). An overview on the general relationships of underlying (exogenous) and proximate factors affecting health and mortality outcomes is presented in Figure 3 which is guided by the frameworks as outlined in Mosley and Chen (1984) and UNICEF (2008). The conceptual core of the framework is the idea that all background variables (cultural, socioeconomic or geographic) have to operate through a limited set of proximate determinants (environmental contaminations, maternal factors, infant feeding habits, and preventive health care practices) which in turn directly influence the risk of disease and the outcome of the disease process.

[insert Figure 3]

The relationship among health inputs and child health outcomes can be written as follows:

$$H_{ijk} = H(E_i^o, P_i^o, v_k, u_{jk}, \varepsilon_{ijk}), \tag{1}$$

where the health of individual i in household j and community k (H_{ijk}) is produced by observed underlying factors (E_i^o), observed proximate factors (P_i^o) and certain unobserved underlying and proximate factors ($v_k, u_{jk}, \varepsilon_{ijk}$).

Further on, equation 2 depicts the mortality production function for individual i, with mortality (M_{ijk}) resulting if health falls below some critical level H^*.[2]

$$M_{ijk} = M(H_{ijk} - H^*) \tag{2}$$

In the empirical analysis, the estimation of parameters for the health and mortality production functions can suffer from three types of problems. First of all, since all health related input variables are treated as exogenous, a bias might arise if we fail to control for simultaneously determined health inputs in the estimation of the health production function. Secondly, health and mortality outcomes are influenced by several individual, family, and community variables. Some of these variables can be observed; others cannot. Table 2 shows how the observed and unobserved factors can be classified in our particular case. The simple association between, for example, the stunting score of a child and the mother's educational level holds if the observed indicator (e.g. educational status) is not correlated with unobserved variables (such as labor market conditions) that affect the stunting score. However, if the unobserved factor affects the child's stunting status and is correlated with the educational attainment of the mother, the estimated effect is biased together with false

[2] Empirically we will model M as the risk of mortality at time t.

standard errors for our parameter estimates. Thirdly, our obtained coefficients can be biased and our standard errors can be false, even if the unobserved factor affects only the outcome variable but is completely uncorrelated with the observed explanatory variables. This might be the case if in the incidence of clustering the mortality risk among siblings and among children residing in the same community is partially due to children sharing the same family and community characteristics. However, the correlation may persist after controlling for observed factors such that the remaining correlation is a consequence of genetic, behavioral, and environmental factors that are common to all children in a particular community or family but that are unobserved. As a consequence, the still correlated observations violate a standard assumption of independence in statistical analyses, resulting in standard errors that are understated and, in the case of non-linear models such as a hazard models, parameter estimates that are both biased and inconsistent (Trussel and Rodriguez, 1990).

To mitigate this problem we adopted 3 strategies. First of all, in order to circumvent the first problem we estimate reduced form regressions and therefore exclude any explanatory variable that we would expect to cause such a problem, which in our case results in dropping the length of breastfeeding from our regression equations. Secondly, we construct and include malaria, health care and a Lake Victoria variable, indicating whether a child lives within 20km around Lake Victoria in order to better capture the specific nutrition and disease environment as explained in more detail in section 4. Thirdly, as explained more in detail in section 5, we use mixed model representations to further address unobserved heterogeneity issues on the family and community level.

[insert Table 2]

4 Geography vs. Ethnicity: The Kenyan Context

The seeming disconnection of anthropometrical indicators on the one hand and health and mortality patterns on the other and the prevailing huge spatial differentials in mortality and undernutrition indicators in Kenya requires a profound investigation of the underlying causes of this phenomenon. Moreover, in order to facilitate the understanding of the country specific context it is further useful to distinguish between underlying causes that have an effect on either child malnutrition or child mortality and those that have an effect on both of them simultaneously. Despite reviewing all of the relevant causes in this section we will only pay attention to those that are particular to the Kenyan context with a main focus on nutritional, epidemiological and cultural factors.

4.1 Nutritional Environment

Food and nutrition availability affects anthropometric and mortality outcomes likewise. While insufficient food and vitamin intake mostly influences child mortality indirectly by increasing the predisposition to diseases it often expresses itself directly in anthropometric measures which therefore frequently serve as proxies of the health and mortality environment in the absence of reliable data on the latter ones. The two most widely used anthropometric indices regarding children are stunting (low height for age) and wasting (low weight for height) which both serve a different purpose. Stunting is claimed to be an indicator of chronic undernutrition resulting of prolonged food deprivation or illness, meanwhile wasting is supposed to reflect acute undernutrition as a result of more recent food deprivation or illness (Nandy et al., 2005).

In addition to this there exist further factors that manifest themselves in both anthropometric indicators very differently. In particular the type of food consumed in the first months of life plays an important role in the growth process of children and hence affects the stunting indicator while the wasting indicator remains relatively unchanged. For the transformation of energy into body growth, certain micro- and macronutrients are very essential like iron, calcium, iodine, vitamin A or proteins (Moradi, 2009). Accordingly, the body on the one hand fails to grow at a normal rate as a result of chronically insufficient intakes of certain types of micro- and macro nutrients and on the other hand exceeds normal growth rates if certain micro- and macronutrients are consumed in high amounts. An important food that makes a vital contribution to the survival, health and body growth of children is fish which provides quality proteins and fats (macronutrients) and vitamins and minerals (micronutrients). Furthermore, it is notable that not fish consumption per-se drives the higher than average growth process. Biomedical research shows that the growth effect due to fish consumption does only seem to occur in combination with a well balanced compositional diet (Marques et al., 2008).

Neglecting the role of diseases on stunting and wasting we would expect that both indicators show high z-values in areas where food availability is high. Moreover, the incidence of stunted children should be particularly low in those regions where fish is widespread available together with other crops and foods.

The Lake Victoria region in fact offers the advantageous conditions just mentioned. Soils are mostly of good quality resulting in agricultural production surpluses what led Fearn (1961) to call the region the 'granary of East Africa'. Moreover, fish is largely available in at least close proximity to the lake despite the strong export orientation of the fishing industry in the context of the nile perch boom. Moreover, the region is situated on the trade route between Tanzania and economically comparatively prosperous Central Kenya which might further help to increase the variability and availability of food in the region.

In general reliable administrative data on soil quality, frequency of rain and agricultural production for Kenya is not available. In the absence of these data it is difficult to asses the food availability and security situation for other regions in the country. Fortunately, USAID created the Famine Early Warning System (FEWS) that issues early warning and vulnerability information on emerging and evolving food security issues in the world. To inform researchers and policy makers FEWS generates the so called Water Requirement Satisfaction Index for maize (WRSI) and the Normalized Difference Vegetation Index (NDVI). Both indices are updated regularly and allow for the investigation of intra-country differences (fews.net). The WRSI for Kenya is used as an indicator of maize performance based on the availability of water to the crop during the growing season. Maize has been selected since it is the most important cereal crop in Sub-Saharan Africa and due to its properties to be cheaper, less water intensive and climatical more robust than other cereals which makes the WRSI an ideal indicator of food security. Looking at the distributional map of the WRSI for Kenya, Figure 4, two things are remarkable. First of all, the areas close to Lake Victoria obtain the highest scores for the whole country indicating a very food secure situation. Secondly, the level of food security deteriorates steadily the further a location is away from Lake Victoria an exception being the coastal area around Mombasa where the food security situation improves again (FEWS-Net, 2004). This result is reinforced by considering the map portrays of the NDVI for Kenya, Figure 5, which is based on meteorological NASA satellites using advanced very high resolution radiometer in order to indicate the vigor and density of vegetation at the earth's surface. An inspection of the map shows a similar pattern with the areas around Lake Victoria showing the highest vegetation density and areas further away depicting continuously decreasing vegetation levels (FEWS-Net, 2004).

[insert Figures 4 & 5]

Relying on the two proxies for food security described above, two main implications can be derived from the previous considerations. Due to a general increase in food security and food availability the closer one gets to Lake Victoria, wasting and stunting indicators should show improvements reaching their lowest values in the area around Lake Victoria. Furthermore, the incidence of stunting should be extraordinary low on the shores of Lake Victoria since fish is widely available as a staple food around this area.

4.2 Epidemiological Factors

Child mortality levels in Kenya declined rapidly after its independence in the early 1960s and reached its minimum levels in the late 1980s. From then on the trend reversed and despite a significant drop in overall fertility rates, mortality levels were continuously increasing up to 115 per 1,000 in the 2003 DHS round (Hill et al., 2004; CBS, 2004). This adverse trend was accompanied by stagnant growth of per capita income, declining levels of immunization, falling school enrollment, and foremost the emergence of the AIDS epidemic (Hill et al., 2004). One of the salient findings of the 1998 and 2003 DHS rounds is the enormous variation of child mortality rates among the provinces reaching 54 per 1,000 in the Central Province and 204 per 1,000 in Nyanza in the 2003 round.

In the following we want to shed some light on the principal underlying causes of the geographic differentials in the observed child mortality rates in order to underscore the role of Lake Victoria on mortality patterns in Kenya.

When looking at the data for Nyanza it surprises that the established link between a good nutritional status, in particular wasting or acute malnutrition, and child mortality risks is not reflected in the health statistics. While children in the Nyanza region show above average scores in the anthropometric indicators, infant and child mortality rates are highest in the region. A first major reason that helps to explain this paradox is the unequal distribution of malaria prevalence within the country - as depicted in Figure 6 - due to regional variations in temperature, humidity, and the existence of bodies of water. Although malaria is epidemic in several areas in Kenya, the Lake Victoria region is the only endemic region in the country with a transmission period that lasts over the whole year (MARA, 2004). Moreover, it is important to note that the risk of malaria infection does not decline in a continuous way starting from Lake Victoria. Due to the elevations of the East African rift valley arising only some kilometers away from Lake Victoria a natural malaria barrier exist in the east that drastically reduces the risk of malaria infection in these regions. In all other regions, except a small stripe on the coastal area, climatical conditions do not favor the reproduction of female anopheles mosquitoes over the whole year due to long periods without rain in these areas. Hence, this restricts malaria transmission rather to the rainy seasons in these areas.

[insert Figure 6]

A second major aspect that affects health and mortality outcomes of young children is the quality of drinking water. Despite being the second largest fresh water lake in the world, the water of Lake Victoria is not safe for drinking and several cases of outbreaks of waterborne diseases are reported each year (Ochumba and Kibaara, 1989; Oguttu et al., 2008; Omwega et al., 2003; Scheren et al., 2000).

The third major difference of the Lake Victoria region compared to other regions in Kenya is the high prevalence of HIV/Aids in the area. While cultural factors, as described later on, can explain part of the high HIV rates in the area, recent studies point to the social erosion of family norms among people around the shores of Lake Victoria. The nile perch boom in the area, starting in the mid 1990s, and the resulting demand for male labor forces in the fishing industry led to a strong influx of migrants into the region which was accompanied by a growing prostitution business (Geheb et al., 2008). Moreover, the increase in demand for male labor shifted the intra-household bargaining power towards men and contributed to weaken the already inferior position of women thereby increasing the likelihood of involuntary risky sexual behavior for women (Béné and Merten, 2008). Further on, Nyanza province is situated on thriving trade and migration routes connecting the economically powerful central area of Kenya with Tanzania. Together with the high urbanization rates in the Nyanza this is likely to contribute to the higher HIV/Aids rates in the area (Oster, 2008).

Taking into account the spatial distribution of mortality drivers as outlined above, we would expect strongly increasing mortality rates in close proximity to Lake Victoria the main reasons being among others the comparatively high HIV/AIDS prevalence and the much stronger predisposition to infectional diseases like Malaria in that area.

4.3 Cultural Factors

Ethnic belonging affects mortality and undernutrition levels in Kenya through a variety of mechanism. Notably geographic and cultural factors in addition to the prevalent political economy play an important role with respect to health outcomes of certain ethnic groups in the country.

Since most ethnic groups in Kenya live spatially concentrated in a very particular region of the country, current administrative provincial boundaries were usually drawn based on the location of a certain ethnicity. The Lake Victoria region is part of the Nyanza province which is predominantly inhabited by the Luo ethnic group. Although Luo have several cultural practices in common with most other ethnic groups in Kenya, there exist three noteworthy differences.

Firstly, while most ethnic groups in Kenya practice male circumcision, Luo besides Turkana, and Itero, which represent only a small part of the population, are known for not being circumcised thereby substantially increasing their risk of HIV infection and HIV related mortality (Chesoni, 2006). Secondly, the type of nutritional intake differs compared to other ethnic groups. Having lived for already more than 400 years in close proximity to Lake Victoria (East African Living Encyclopedia), Luo have benefited from the beneficial food availability and protein situation, therefore showing significantly better mean height values for men in historical data compared to all other Kenyan ethnic groups (Moradi, 2009). Given the long time span of settlement close to the Lake Victoria and the historical data, one might speculate over emerging genetic differences among the Luo and other ethnic groups that lead to mean height advantages that manifest themselves already in nutritional indicators in early ages. Following the recent WHO study on the new child growth standard genetic factors seem to play a minor role in explaining the disparities in physical growth among children (WHO, 2006).[3] Thirdly, fish consumption in Kenya is not only determined by availability aspects but as well by cultural habits. While Luo and some ethnic groups in the coastal area use fish as a staple food, it is viewed with considerable suspicion among ethnic groups in Central and Eastern Kenya (Oniang'o and Komokoti, 1999; Peters and Niemeijer, 1987).

With respect to differences in the position of women among Kenyan ethnic groups, no clear picture exists. While polygamy is common among most of Kenya's ethnic groups widow inheritance is practiced primarily by the Luo and certain smaller clans among the Luhya ethnic group, therefore weakening the role of women in these ethnic communities. In contrast, female genital mutilation is practiced by the majority of Kenya's ethnic communities, while only Luo, Turkana, Luhyia and Iteso do not (Chesoni, 2006). Moreover, female education levels tend to be high among Luo women compared to other ethnic groups (Wainaina, 2006).

Furthermore, ethnic belonging plays a crucial role on the allocation process of

[3] To further investigate this issue, we compared children's mean stunting scores between Luo and other ethnicities outside the Lake Victoria Region and additionally outside Nyanza and Western province. Based on an oneway ANOVA, differences in height for age scores turned out to be statistically insignificant for this setting. Therefore, we conclude that genetic differences do not explain the observed growth differential for children in our context.

public resources and political positions in Kenya due to the prevailing kinship structures and patron-client relationships (Cohen, 1995; Miguel, 2004; Miguel and Gugerty, 2005; Weinreb, 2001) and this way affects health indicators. Out of the more than 40 ethnic groups in Kenya, the Luo represent about 13% of the Kenyan population and constitute the third largest ethnic group in Kenya whereby only the Kikuyu with 23% and the Luhya with about 14% tend to have higher shares in the overall population.

Although Luo ethnic groups took an important role in the independence process in Kenya, they have been politically under-represented at national political levels and except very recently not being part of any coalition since 1965. The Luo are the only major ethnic group in Kenya that has not been part of the national government since this time span and this under-representation of Luo interests on the national level has resulted in a limited access to public funds from the national level and lead to a steady under-investment of health and schooling facilities in the Nyanza region (Alwy and Schech, 2004; Muhula, 2008; Nyanjom, 2006) compared to most other regions, the exception being the north east of Kenya.

Bearing in mind the circumcision behavior of ethnic groups in Kenya, we would expect the highest HIV/AIDS prevalence among the Luo ethnic group. In addition, the situation of the political economy in Kenya is likely to further aggravate the mortality levels of the Luo due to a worse access to health care facilities compared to the main other ethnic groups in Kenya (Cutler et al., 2006). Since discriminatory practices in the allocation process of public resources probably occurs in practice on a provincial level meaning that relatively less money will go to Nyanza as a whole, there might be an unfavorable effect on mortality levels for all ethnic groups living in Nyanza.

5 Empirical Findings

5.1 Data

The KDHS 2003

In the empirical analysis we use data from the 2003 round of the Kenyan Demographic and Health Survey (KDHS). The KDHS 2003 includes full birth history information from 4346 women of reproductive age that gave birth to at least one child in the five years preceding the 2003 KDHS survey. For the first time the KDHS includes data from all provinces of Kenya as well as data on HIV testing. Moreover, the survey is based on a two-stage survey design. In the first stage 400 clusters were randomly chosen from a master frame. Afterwards, households were systematically sampled out of each cluster.[4] In every second household sampled, men, aged 15 to 54 years, were interviewed to conduct a Men's questionnaire. All women and men living in households selected for the Men's questionnaire were asked to voluntarily participate in the HIV testing.

[4]In the following we often refer to clusters as communities since in the DHS context it is a geographical unit, consisting of several households.

76% of all eligible women voluntarily agreed to undergo the test.[5]

In addition to the variables directly derived from the household questionnaires, we calculate the distance of each cluster to the shores of Lake Victoria using the GPS coordinates provided by ORC Macro. We define the Lake Victoria region as the area within a 20km boundary to the shores of the lake. Furthermore, we exploit the MARA (2004) database on endemic malaria to obtain district level information on malaria prevalence in Kenya.[6] Unfortunately, we could not obtain data on the health care sector in Kenya on the district level. Instead, we rely on data published in Nyanjom (2006) who reports information on the number of people per medical officer and public health expenditures per capita on provincial level for the time period 1995 - 1998.

The samples

As is common in the literature we include only those children who were born within the 5 years preceding the survey. Since hygienic and socio-economic conditions are less likely to have changed over the course of 5 years compared to 10 or more years, this decision improves the accuracy of the matching of the covariates to the actual survival time in the multivariate analysis of under-5 mortality. Moreover, data on children's height and weight has only been collected for those children being below the age of 5 at the time of the survey therefore restricting the information on the nutritional status of children to the same period of time.

The mortality sample consists of 1368 mothers reporting 2697 births in the last five years. 605 of these children died. The respective undernutrition sample remains with 1218 (1217) mothers who reported data on 1704 (1701) living children at the time of the survey for the stunting (wasting) regressions.[7]

Variables of interest

The selection of variables for descriptive statistics and the undernutrition and mortality regressions is guided by the frameworks outlined in the previous section and the discussion of the role of ethnical, political and geographical factors in section 4 of this paper. An overview of variables used in this article including its coding is provided in Table 3.

[insert Table 3]

Two variables deserve particular attention. First of all, as mentioned earlier, we take the HIV status of the mother into account. The interpretation of

[5]The official KDHS 2003 report provides several descriptive and multivariate examinations on whether non-participation in HIV testing is systematically related to other variables. No systematic relationship was found (CBS, 2004) and therefore we expect our results not to be effected by sample selection bias when using the reduced HIV sample.

[6]The same data base from http://www.mara.org.za/ was used in Oster (2007) to calculate regional malaria prevalence rates. Moreover, at the country level these malaria measures are closely correlated with climate-determined malaria susceptibility, as used in Sachs and Malaney (2002).

[7]Data on childrens' height and weight was missing for 11% of all living children.

this variable in the regression context is not straightforward. Since the HIV status of children has not been collected, it remains unclear whether the AIDS virus has been transmitted to the child at all during pregnancy or breastfeeding period and whether the mother already had the virus at the time of the birth of the child. Incorporating the HIV status of the mother in the regressions is therefore likely to yield a downward biased coefficient with a lower significance level. Moreover, the HIV status of the mother does not only measure a direct epidemiological effect on children but as well a socioeconomic one. In particular, children in a HIV affected household might suffer from diminishing capacities of their main caregivers to purchase certain key inputs for the children due to a loss of household income as a result from the disease. Furthermore, as described in section 4, the HIV status of a parent is partly related to cultural practices, e.g. male circumcision and therefore inhibits cultural elements as well. Secondly, the distance of a cluster to Lake Victoria plays an important role in our study. As pointed out in section 4 we would expect to see much better stunting values in close proximity to Lake Victoria due to the availability of fish and other food over the whole period of the year while with respect to wasting we would assume constantly improving wasting rates the closer Lake Victoria. In contrast, we would infer under-5 mortality levels to substantially deteriorate in close proximity to Lake Victoria due to the much higher disease environment in this area. Since most of our health environmental and geographical variables are either only on provincial or district level and moreover might not be free of measurement error, we would still expect to have an effect on our distance variables. In order to measure these effects appropriately we include a dummy variable indicating whether a household lives in the Lake Victoria region, within a distance of 20km of Lake Victoria, in the stunting and under-5 mortality regressions while in the wasting regression the distance to Lake Victoria is incorporated as a continuous variable.

From the economic literature (Mosley and Chen, 1984; Smith and Haddad, 2002) as well as from our theoretical framework, it becomes clear that the variables described in Table 3 are important to study the determinants of undernutrition as well as the context of under-5 mortality. Thus, we use the same list of covariates in the multivariate analysis of undernutrition and mortality.[8] The KDHS includes some further variables, e.g. information on children's protein intake or pre- and post natal care, which are only used for descriptive purposes since these variables exhibit a very large number of missing observations leading to a too strong reduction in the sample size for the multivariate analysis.

5.2 Descriptive Statistics

Summary statistics on the variables used in this study are provided in Table 4. Moreover, we distinguish in Table 4 between different geographical and ethnical specifications. Column 1 depicts statistics based on the Lake Victoria region,

[8]The final model specifications include squared terms whenever the respective coefficient showed a statistically significant value. Otherwise squared terms were excluded. In this regard model specifications might differ between the stunting, wasting and mortality regressions.

the area within a 20km distance of the shores of the lake, while column 2 provides information based on overall Kenya except the Lake Victoria region. Column 3 and 4 refer exclusively to the Lake Victoria region. In column 3 summary statistics are provided for the Luo ethnicity while statistics on the remaining ethnic groups in the area are shown in column 4. In addition, columns 5-12 comprise information on the same set of variables for overall Kenya and for each of the eight Kenyan provinces.

The first 3 lines of Table 4 already demonstrate the distinct setting of the Lake Victoria region in Kenya in terms of child malnutrition and mortality outcomes. As discussed in section 4, average rates in stunting and wasting are far below average in this region. Quite the contrary, under-5 mortality is by far the highest in Nyanza, reaching its peak in the Lake Victoria region. Bearing in mind that the underlying causes of malnutrition and mortality may differ, we look separately at geographical and ethnic disparities in the variables of interest.

[insert Table 4]

Malnutrition

The results from Table 4 show that the dietary intake is much higher in the Lake Victoria region than in the rest of the country, suggesting higher food availability at the shores of Lake Victoria. The observed higher than average intake of local grains and vitamin A rich fruits, like mango or papaya, can be attributed to the fertile soil found in the lake basin in combination with enough rainfall, facilitating a large supply of these aliments. Moreover, 38% of Luo mothers in the Lake Victoria region allocate protein rich food to their children at least once a week, compared to only 26% of mothers belonging to other ethnic groups in the region and to a national average of 22%. This remarkable difference between Luo and other ethnic groups within the Lake Victoria region might partly be explained by different food preferences as outlined in section 4.3 whereby other ethnic groups do not use fish as a staple crop. Figures 4 and 5 suggest more favorable agricultural conditions the closer the lake (FEWS-Net, 2004), substantiating the finding of highly cultivable soil near Lake Victoria. In the DHS, data on protein intake is not further disaggregated into its share of fish, meat or eggs. We use secondary data to stress the relevance and availability of fish. Data from the Kenya Integrated Household Budget Survey (KIHBS) 2005/2006 shows that fish consumption is highest in Nyanza province. In addition, households in Nyanza seem to spend 6.1% of their budget for food on fish compared to a national average of 2.1% (Kenya National Bureau of Statistics, 2006).

Mortality

The high mortality rates in the Lake Victoria region point to the existence of extraordinary factors that help to explain the observed outcomes.

Table 4 depicts considerably higher malaria, HIV and diarrhea prevalence in the Lake Victoria region than in all other parts of Kenya. This result confirms the findings from the literature review in section section 4. Moreover, high malaria rates are not only confined to the Lake Victoria region and Nyanza but to a lesser extent as well to Western province which might be indebted to some of its area bordering Lake Victoria. Furthermore, the incidence of diarrhea in the Lake Victoria region (21.7%) is clearly above the Kenyan average and increases about 4% compared to the Nyanza average which might indicate the poor quality of drinking water of Lake Victoria and in its connected open waters. Further on, HIV rates in Nyanza and in particular in the Lake Victoria region are much higher than in any other area of the country which seems to be partly due to its comparatively high level of urbanization and its extraordinary position as a traffic hub between Tanzania and Central Kenya.

Geography plays an important role on under-5 mortality. Ethnicity may be as important. The Lake Victoria region is – besides its unfavorable disease environment – characterized by being predominantly populated by the Luo ethnic group. The Luo represent 81% of the total Lake Victoria population and 94% of the Lake Victoria population in the province of Nyanza. The small upper northern part of the Kenyan Lake Victoria region belongs to Western province and is mainly populated by the Luhya ethnic group who also represent almost the entire remaining population of Nyanza (17%). Distinguishing between ethnic groups within the Lake Victoria region, enables us to disentangle the effect of geographical and cultural factors. The Luo exhibit significantly worse outcomes in several proximate factors of child mortality than the other ethnic groups around Lake Victoria. 26% of all Luo mothers are tested HIV positive. Compared to a national average of 8.8% and to an average of 10.6% for the remaining ethnic groups in the Lake Victoria region, the result clearly points at a strong relationship between cultural habitus of the Luo and HIV infection, as already described in section 4. Moreover, the size of the differences between ethnic groups in the Lake Victoria region astonishes. While circumcision practices have been pointed out as one of the potential reasons for higher HIV rates among the Luo, the difference in most DHS reports for SSA countries between circumcised and uncircumcised adults was substantially lower and usually amounts to 4% - 7%.

A similar picture emerges from maternal factors and the pre and post natal behavior of the Luo. Luo mothers seem to start much earlier with bearing children than Luhya and most other ethnic groups. In addition, birth intervals between consecutive children are considerably shorter reflecting higher total fertility rates among Luo than among any other ethnicity (CBS, 2004). Moreover, short preceding birth intervals are often caused by the death of the previous child. Short succeeding birth intervals result in termination of breastfeeding. Indeed, both indicators, previous dead child and breastfeeding, are found to be especially under performing for Luo. Furthermore, the average number of child deliveries and caesarean sections as well as pre-birth visits in official health centers all show the lowest value for Luo. Besides of cultural habits of the Luo, these outcomes may also point to discriminatory practices against the Luo from the national level resulting in limited access to public funds and hence to lower public health expenditures. Such an interpretation is supported by the data on the Kenyan health care sector obtained from Nyanjom (2006). This data indicates that Nyanza receives the lowest amount of public health expenditures per capita and further on has the highest ratio of inhabitants per physician among all Kenyan provinces.

The descriptive findings are in line with the considerations undertaken in section 4. Thus, geographical, cultural and political factors seem to contribute jointly to the high mortality rates in the Lake Victoria region compared to the rest of Kenya. Due to the simultaneous occurrence of all three of these factors in the Lake Victoria region, it is difficult to establish the influence of a certain factor on our anthropometric and mortality outcomes when relying on bivariate statistics and analysis. Therefore in the following section, we will use regression analysis to examine causal relationships going from the observed covariates to the malnutrition and mortality outcomes, thereby putting a special emphasis on geographic and cultural factors.

5.3 Method

In order to investigate the determinants of undernutrition we rely on a linear regression model, while in the context of under-5 mortality we use a (non linear) Cox proportional hazard model. We will use *multilevel* extensions of the respective models for various reasons. Multilevel modeling allows for efficient and, for non linear models, consistent estimation in the case of significant intragroup clustering. Beyond, we will exploit the variance decomposition, inherent in multilevel modeling, to assess the contribution of covariates on the within-level variation.

The survey design of the KDHS 2003 involves hierarchical collection of data at the family and community level which results in clustering of undernutrition and mortality outcomes. In regression analysis, clustering is problematic if it is not only due to observed but also to unobserved household and community factors. For instance, net of observed factors, people within the same community are more alike than people across other communities since they are likely to share similar latent characteristics. Therefore, the statistical assumption of independence of the error term is violated and as a consequence confidence intervals are underestimated leading to false statistical inference. Typical unobserved factors at the family level are shared genetic factors, social practices or the pre and post natal behavior of the mother (Bolstad and Manda, 2001). Likewise, shared environmental factors may lead to clustering at the community level.[9] To illustrate the extent of clustering in the KDHS 2003 data, we focus on the observed under-5 mortality outcomes. In our sample for all of Kenya 2915 (67%) of the 4346 interviewed women did not experience any child deaths, while 941 (21.7%) women had to suffer from the death of exactly one child. Just 184 (4.2%) women experienced three or more child deaths in the five years preceding the survey. These 4.2% account for more than 30% of all deaths, showing a substantial amount of correlated outcomes and clustering

[9] See section 3 or Sastry (1997a) for a detailed overview over potential unobserved family and community characteristics.

within families. A similar pattern arises at the community level. 62% of the 400 clusters under consideration contribute to 26% of all dead children, while on the other hand 21% of the communities account for more than 50% of all child deaths.

Due to the large clustering of outcomes on the mother as well as on the community level, we use multilevel mixed effects variations in the respective regression models. We use three-level models, controlling for correlated outcomes among siblings, i.e. on the household level, and among communities.[10] In these models the error term is decomposed into a single error term on each level, capturing unobserved heterogeneity at each level. Thereby, multilevel models entail another advantage to classical linear regression models. We will especially exploit the option of variance decomposition to measure the explanatory power of the covariates on the between family and community variation and, even more, to shed light on the specific contribution of each variable on the overall variation of malnutrition and mortality outcomes across families.

5.3.1 The linear multilevel model for malnutrition

To analyze the driving factors of malnutrition, we use a linear multilevel regression model. As mentioned above multilevel models are applied to control for clustering caused by unobserved heterogeneity and to increase the precision of the estimated coefficients of the covariates.[11]

A linear multilevel model can formally best be described by a two level model with only one explanatory variable. The idea of the model, also called random intercept model, is that the intercept of all individuals is not identical but group specific. Each group, e.g. family, has a specific effect on the outcome variable.

The model equation reads:

$$Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + \varepsilon_{ij}, \tag{3}$$

where Y is stunting and wasting, respectively and X is a vector of observed covariates.

The regression coefficient β_{0j} can be expressed as:

$$\beta_{0j} = \gamma_{00} + U_{0j} \tag{4}$$

Finally, the combined model can be expressed as consisting of a fixed part (first term) and a random part (second term):

$$Y_{ij} = (\gamma_{00} + \beta_1 X_{ij}) + (U_{0j} + \varepsilon_{ij}) \tag{5}$$

The random effect for each family is not measured itself. Its variation across families is estimated under the assumption of its independence with the regressors and being normally distributed with zero mean and variance θ. The

[10] Using a two-level model by neglecting correlations on the community level would lead to an overestimation of the family random effect.

[11] Goldstein (2003) and Hox (1995) give a comprehensive overview of the statistical theory underlying multilevel modeling. For a good overview of applications in this area, see DiPrete and Forristal (1994) and Guo and Zhao (2000).

random effect captures unobserved group specific effects. If undernutrition outcomes were independent of families – outcomes of siblings controlling for observed effects would be uncorrelated – the variation of the family random effect would be insignificant.

Since we control for family as well as for community specific effects, our three level random intercept model reads:

$$y_{ijk} = (\gamma + \beta_1 x_{ijk}) + (v_k + u_{jk} + \varepsilon_{ijk}) \tag{6}$$

with $i = 1, ..., I$ individuals, $j = 1, ..., J$ households and $k = 1, ..., K$ communities. u_{jk} is the household random effect, v_k the community random effect.

The models are estimated using the "xtmixed" command implemented in Stata (Stata, 2007).

5.3.2 The multilevel Cox frailty model for under-5 mortality

Cox proportional hazard models, proposed by Cox (1972); Cox and Oakes (1984), are the standard models used in child mortality analysis (Cameron and Trivedi, 2005). In a proportional hazard model, the hazard rate is the instantaneous risk of death in t conditional on survival up to t:

$$\lambda(t|x) = \lambda_0(t) exp(x'\beta). \tag{7}$$

λ_0 is the baseline hazard only depending on t. On the contrary, $exp(x'\beta)$ depends only on x. The idea is that all hazard functions are proportional to the baseline hazard, just being shifted by the scale factor $exp(x'\beta)$. $\lambda(t|x)$ is the hazard of child death at time t given x. The advantage of the semi-parametric Cox proportional hazard model over parametric models is that no functional form of the underlying hazard function has to be assumed.

The multivariate kindred frailty model has been developed by Vaupel (1989, 1990). The two-level model was applied to study the effect of unobserved shared family characteristics on survival. Sastry (1997b) and later on Bolstad and Manda (2001) in a full Bayesian approach and Pankratz et al. (2005) extended the model to the multilevel case.[12] In a standard frailty model, a frailty, z, is an unobserved random effect which works multiplicatively on the hazard function.

$$h(t|x, z) = z\lambda(t|x). \tag{8}$$

In economic survival analysis, the frailty is usually referred to as a shared frailty since it is a random effect which is the same for all members of a group, for example a family or a community effect (Anderson et al., 2007). Transferred to the three level model with unobserved frailty on the household and community level, the hazard function reads:

$$h_{ijk}(t|x_{ijk}, u_{jk}, v_k) = u_{jk} v_k \lambda(t|x_{ijk}), \tag{9}$$

with individuals $i = 1, ..., I$, $j = 1, ..., J$ households and $k = 1, ..., K$ communities. u_{jk} is the household random effect, v_k the community random effect.

[12] Guo and Zhao (2000) give a good overview over multilevel modeling for binary data.

The individual (child) frailty is absorbed in the baseline hazard. The unobserved frailty is assumed to be independently distributed of all covariates and to follow a Gamma distribution with mean 1 and variance θ (Sastry, 1997b).

The family frailty effect measures the variation in family specific exposure to risk across families after controlling for observed variables. Children of families with a large frailty have, ceteris paribus, a larger risk of dying. This excess risk would be triggered by unobserved behavioral or genetic family specific factors. Insignificant variation of the family frailty effect would mean that there are no unobserved family specific characteristics. In this case, survival chances of siblings would be uncorrelated.

The model is estimated using the "coxme" command of the kinship package in R (Therneau, 2006).

5.4 Regression Results

In all of our regressions, we start with a conventional single level regression specification, called Model I or the 'Standard Model'. In Model II a family random (frailty) effect is incorporated into the model. Finally, we add in Model III a community random (frailty) effect. Likelihood ratio tests are applied to test for significant random effects in the linear as well as in the non linear model (Goldstein, 2003). Since Model III performs - based on the likelihood ratio tests - significantly better than model II and I for all of our analyses, we will discuss our regression results based on this model. We will, however, compare the results of Model III with the inferior models to justify multilevel modeling.

5.4.1 Undernutrition

Results of stunting and wasting regressions are shown in Table 5 and 6, respectively.

[insert Table 5]

Stunting

In Model III, stunting values are significantly better for girls than for boys which has been a common finding in the literature on SSA countries (Svedberg, 1990; Klasen, 1996; Wamani et al., 2007). Being a twin shows a significant negative effect which seems reasonable given the higher nutrition competition and the circumstance that twins are smaller and lighter at birth compared to single births. Children of better nourished mothers, as indicated by a higher body mass index, have, ceteris paribus, a higher height for age score. Moreover, the higher wealth, educational attainment of the mother, and the better the access to water, the better the stunting values of a child. On the other hand, the age of women at birth as well as living in rural areas, or having access to better sanitation facilities does not seem to have an effect.[13][14]

Interestingly, a child's disease status (hiv, diarrhea, malaria) does not have a significant impact on a child's height for age status. Diseases have a rather short-term impact on children's health outcomes. The long-term indicator stunting may therefore fail in displaying any significant negative relationship.

Children residing in close distance to Lake Victoria are showing significantly more favorable height for age outcomes. As pointed out in sections 2 & 5.2 the high availability of fish in combination with very advantageous agricultural conditions are likely to be the main reason for this puzzling result and the extraordinary low prevalence of stunted children in the region.

These results endorse, once again, our afore mentioned theoretical considerations. These results endorse, once again, our afore mentioned theoretical considerations. Adverse geographic factors in the Lake Victoria region - malaria suitability, HIV endemicity and possibly unsafe drinking water - may affect the excessive risk of death for children in the region. They do, however, not affect a child's stunting status. Quite the opposite, the region around Lake Victoria exerts positive influences on children's nutritional outcomes. Fertile soils, a high level of food security and high protein availability (fish) foster the growth process of children. This result does not question the clear epidemiological relationship between a child's disease status and its mortality outcome. But the findings challenge the adequacy of inferring children's health and mortality outcomes from children's height status. Environmental factors - as found in the Lake Victoria region - can substantially modify the relationship between children's growth and health status.

Wasting

In contrast to stunting, wasting outcomes of children are much stronger affected by short term factors and, hence, are subject to much larger fluctuations. A poor wasting status is often a consequence of having *recently* suffered from illness or from insufficient food intake (Fawzi et al., 1997; UNICEF, 2008). As a consequence of these short term variations and missing information on important factors wasting regressions are usually not considered by demographers and economists. Technically speaking, a lot of the variation in wasting indicators among children is due to short-term variation, e.g. whether a child got sick or not, which can not be captured appropriately by the set of variables obtained from common household surveys. Therefore, determinants of anthropometric outcomes in wasting regressions are much more difficult to establish than in stunting regressions. Nonetheless, section 4.1 provides a priori reasons to believe that there are very crucial geographical factors at play that should be able

[13]The significant positive coefficient of the variable "people per physician" is most likely driven by the tremendous share of people per physician in the North Eastern province which exhibits, however, good stunting outcomes.

[14]Once again, standard errors of all estimates mentioned above slightly increase from Model I to Model III. The significance of the variables, first born child, birth interval and Luo vanishes when controlling for unobserved heterogeneity in the multilevel specifications. Neglecting highly correlated outcomes of children within families and communities leads to false statistical inference, justifying the usage of multilevel models.

to be verified even in the multivariate case. In particular, the observation that food insecurity constantly increases throughout the country, the further away Lake Victoria, demands to be investigated. Instead of using a dummy variable referring to the Lake Victoria region, we introduce a continues variable measuring the distance to the Lake Victoria in km. Regression results are presented in Table 6.

[insert Table 6]

As expected, most of the variables fail in explaining wasting differentials. In our final model, Model III, coefficients of mother's educational status, the water supply index and the first child dummy have the expected sign but are slightly insignificant. Just four variables appear significant. Mother's BMI, itself an indicator of nutritional status, is positively affiliated with children's weight. The quality of the sanitation facility effects positively the weight-for-age status.

Once again, geographic and political factors seem to play a major role in determining nutritional and anthropometric outcomes. The lower a province's ratio of people per physician, i.e. the easier the access to medical care in that province, the higher the children's wasting score. The closer Lake Victoria, the higher average z-values of wasting.[15]

5.4.2 Mortality

Results of the under-5 mortality regressions are presented in Table 7.[16] Once again, Model I is the standard single level regression model. Model II incorporates a household level random effect. In Model III a community effect is included in addition to the household effect. Regression results will again be reported based on the statistically superior model, Model III.

[insert Table 7]

In Model III, coefficients on all variables that are commonly included in under-5 mortality regressions show a reasonable economic size, take the expected signs and are in line with the empirical literature. The hazard rate is approximately 20% lower for girls than for boys. Being a twin increases the probability of death by 50%. The location of the family, in terms of rural or urban, does not have a significant impact on survival chances. Interestingly, the same is true for the body mass index of the mother. Children stemming from better nourished mothers do not exhibit a lower risk of death. Children from richer households do, ceteris paribus, not seem to have a lower probability of

[15] As explained in section 4.1 we would expect the distance effect to vary continuously with respect to wasting outcomes. Robustness checks which are not reported here confirm that distance to Lake Victoria only seems to matter when captured continuously. In the full model the distance effect when included as a dummy variable always turned out to be statistically insignificant.

[16] Hazard rates, the probability of death in t conditional on survival up to t, are to be interpreted in relation to 1. Thus, a hazard ratio of 1.2 implies a 20% higher risk of death.

death.[17] The educational background of a mother evolves as highly significant, showing that children's survival chances are higher, the better educated the mother. Neither the quality of the water source nor the hygienic status of the household seems to have a considerable impact on under-5 mortality in Kenya. Moreover, a higher age of the mother at birth, a longer preceding birth interval and being the first born child in a family, increases a child's survival probability. Significant squared terms with reversed signs implicate diminishing returns of the factors.

Before examining the context specific determinants of under-5 mortality, it is useful to briefly discuss the benefits obtained from the adopted multilevel modeling strategy in this particular case. Comparing the highly significant family random effect of 0.60 in Model II with the family random effect of the unreported regression model without covariates but only the family random effect, of 0.90, allows to conclude that the observed covariates were capable of explaining approximately 33% of the variation in child deaths across families. The family random effect of 0.60 is, however, likely to pick up unobserved heterogeneity on the community level; meaning that variation in child mortality across families could partly be explained by disparities in unobserved characteristics across communities. Indeed, the family random effect decreases by another 30% to 0.42 in Model III. The significant unobserved heterogeneity on the community level of 0.16 might be due to differences in unobserved child or family factors across communities. Since we included a large set of variables on both levels, we are, however, interpreting this effect as stemming from unobserved community, i.e. geographic or infrastructure, characteristics, that play a major role in explaining differences in mortality rates across Kenya. As expected, after the inclusion of family and community random effects standard errors of almost all covariates increase, lowering their statistical significance. Moreover, we do not only observe changes in the standard errors, but as well differences when comparing the hazard estimates of the models. Consistent with previous literature on multilevel frailty models (Omariba et al., 2007; Sastry, 1997a), effects of socioeconomic household level variables (assets, education) increase, while individual risk variables (birth order, previous child died) decrease. Especially striking is the decreasing effect of the variable that indicates whether the previously born child is still alive. This variable, highly correlated among siblings, is a clear indictor of families with higher mortality risk. The higher the unobserved frailty effect, the higher the probability that a sibling has already died. Neglecting the frailty leads to an overestimation of this effect. Contrary, factors independent of the unobserved frailty and among siblings, e.g the sex of the child, or the age of the mother, show only slight changes.[18]

Taking into account the specific cultural, geographic and political context we included variables on ethnicity, HIV-status, malaria and diarrhea prevalence, the Lake Victoria region, the number of people per physician and per capita

[17] This missing significance might to some extent result from the high correlation (0.47) of the asset index with the educational status of the mother.

[18] Underestimation of socioeconomic variables is a standard result in hazard models that are not controlling for unobserved heterogeneity. The mathematical foundation can be in found Guo and Rodriguez (1992)

health expenditures. Given the remarkable correlation of these factors, a simultaneous inclusion seems crucial to get unbiased estimates. All of these variables turn out to be significant.

Even after the inclusion of behavioral characteristics, being born to a Luo mother seems to have an adverse effect on survival chances. Controlling for geographical factors and the HIV-status, this result strongly points to an influence of adverse cultural practices (pre and post natal behavior), which were either unobserved or could not be considered in the regression. The potential influence of the latter has been illustrated in section 5.2 by the low share of Luo bearing their child in official health centers and the low vaccination rates of Luo's children signifying the awkward pre and post natal behavior of Luo. Omariba et al. (2007) undergo a similar study on child mortality in Kenya neglecting, however, HIV-infection and geographic factors. They conclude an excess risk of 10.5 for Luo children compared to children from Kikuyu[19] mothers. Their much higher coefficient suggests an exaggeration of the Luo effect based on missing but highly correlated variables in their study.

The HIV status of the mother has severe negative impact on the survival status of a child. We estimate a hazard rate of around 1.5 bearing in mind that the coefficient is likely to be underestimated.[20]

Based on the results in Model III, the excess risk of living in a high risk malaria area is around 35% and therefore we confirm that malaria prevalence has a strong impact on mortality rates. Since to our knowledge no studies exist that tried to explicitly capture malaria prevalence as a covariate in a comparable setting we are not able to compare the magnitude of the coefficient to other studies directly. Nonetheless, recent studies from the medical literature (Snow et al., 1998; Omumbo et al., 2004; Ndugwa et al., 2008) commonly underscore the still massive impact of malaria on mortality in Kenya and Sub-Saharan Africa as well.

Besides malaria and HIV, diarrhea is presumed to be one of the most important drivers of child mortality. Our estimation suggests that living in a high risk diarrhea community almost doubles the excess risk of dying. This effect is possibly triggered by poor water quality. Even though we controlled for the type of water source, no direct indicator of water quality was available.

The negative effects of limited access to public funds are eminent. Children residing in provinces that receive lower public health expenditures and suffer under a higher share of people per physician exhibit significantly higher risk exposure. Political discrimination seems to be an important factor of the spatial variation in under-5 mortality rates.

Despite controlling for geography, ethnicity and political outcomes, we still obtain a positive and significant coefficient for the Lake Victoria region on

[19]Living primarily in Nairobi, Central province and the Rift Valley, the Kikuyu represent the largest ethnic group in Kenya.

[20]In a longitudinal survey for rural Tanzania, Ng'weshemi et al. (2003) report a child death hazard ratio for maternal HIV infection of 2.3. Zaba et al. (2005) find in a cohort study for Uganda, Tanzania, and Malawi a hazard rate of 2.9. In a retrospective panel data study for Burkina Faso, Becher et al. (2004) estimate the hazard rate of mother's survival status to be 5.4 for children aged 1-5.

under-5 mortality. This demonstrates the unique misanthropic environment of the region, which could not be sufficiently captured by the variables at hand.[21] The positive dummy might also pick up the natural malaria barrier of the East African rift valley - arising only some kilometers away from Lake Victoria - reducing the risk of malaria drastically.

The regression results confirm the theoretical considerations and descriptive findings of sections 4 and 5.2, respectively. The vast mortality rates of the Lake Victoria region rest upon a *simultaneous* impact of unfavorable geographic, cultural and political factors. There might be other regions in Sub Saharan Africa showing severe water pollution, elevated climate suitability for malaria transmission and high susceptibility for HIV infection. In the case of the Lake Victoria region of Kenya, these adverse geographic conditions are, however, found in a territory which is not only primarily populated by an ethnicity demonstrating adverse pre and post natal behavior but which is also suffering from political discrimination leading to underdeveloped access to health infrastructure.

5.4.3 Family clustering decomposition

The regression analyses conducted above provided insights on causal effects of potential individual malnutrition and mortality drivers. From a policy perspective it is of interest to know those covariates that explain most of the variation in malnutrition and mortality outcomes across families or communities. Which covariates link children's risk of malnutrition or children's survival chances within a family or community? Gaining knowledge of these key variables allows policy makers to improve the targeting performance of poverty alleviation programmes.

An insight to this question can be derived by observing the variation of the family or community random effect (frailty) when each covariate is included exclusively in the null model or omitted exclusively from the full model. If the specific variable accounts for a large part of the overall variation, the variance of the random effect should decrease substantially when included in the null model and increase substantially when omitted from the full model. We will conduct this test for each variable for the stunting, wasting and mortality regression. To keep things clear we will limit this analysis for the clustering of malnutrition and mortality outcomes within families. Therefore we will use the Model II specification.

The results are shown in Table 8. The first line depicts the variation of the family random effect when no covariate is included (null model), and when all covariates are included (full model). The results indicate that the entire set of covariates accounts for 27%, 39% and 34% of the amount of family clustering, i.e. the between family variation, for the stunting, wasting and mortality regression, respectively.

Concerning the stunting regression, a mother's HIV and children's diarrhea status account for the largest proportion of the between family variation in

[21] Recall, that the risk exposure to diarrhea and malaria was measured on the community and district level, respectively. Moreover, data on public health expenditures was available on the provincial level only.

stunting outcomes. This result holds for the inclusion of the variables in the null model and their exclusion from the full model. The same variables appear to be largely responsible for the family level clustering of wasting outcomes. The BMI of the mother emerges as another important factor for the between family variation in wasting rates. As for the regression results, a different pattern emerges when considering the under-5 mortality within family clustering. Mother's education, the survival status of the previous child and being the child of a Luo mother matters the most when explaining differences in mortality outcomes between families in Kenya.

These results should not be confounded with the regression results of Tables 5 - 7. For example, being a female or living around Lake Victoria has a significant impact on a child's nutritional and mortality outcome. Those variables are, however, better suited to explain the between sibling and between region variation respectively, than the between family variation.

[insert Table 8]

6 Conclusion

Kenya's Lake Victoria region is marked by an interesting puzzle. Under-5 mortality rates are by far the highest in the country, while at the same time anthropometric indicators of children show remarkable good values. The extent of this abnormity becomes even more astonishing when comparing the Lake Victoria area to other regions in Sub-Saharan Africa. Nowhere else in the whole of SSA we find such a strong disconnection of anthropometric and mortality outcomes.

In order to examine and understand the causes of this unusual phenomenon we undertake the uncommon step to analyze the determinants of mortality and undernutrition jointly. Moreover, to reduce the likelihood of obtaining biased and inefficient estimates in our multivariate regressions we construct a new set of context specific variables that supplements the conventional DHS data in addition to the application of suitable multilevel modeling techniques.

Our findings point to a unique interplay of cultural, geographical and political factors in the Lake Victoria region which are responsible for causing the described paradox. Concerning the under-5 mortality pattern in Kenya and around Lake Victoria we find that a salient disease environment characterized by extremely high malaria prevalence, polluted water sources and high rates of infectious diseases like HIV/AIDS is one of the key drivers of the massive under-5 mortality rates in the lake region. Furthermore, we see that even after controlling for mother's age at birth, birth spacing, birth order and HIV-status an ethnic specific effect remains. Being born to a Luo mother affects survival chances adversely, most likely based upon unfavorable unobserved pre and post natal behavior. Political discrimination does also seem to be an important factor of the spatial variation in under-5 mortality rates. Children residing in provinces that receive lower public health expenditures - such as Nyanza province - exhibit significantly higher risk exposure. In addition, the results indicate that even after inclusion of a rich set of covariates and controlling for clustering in unobserved characteristics, we are still confronted with an unusual high mortality rate in the Lake Victoria region that remains unexplained by the covariates and that is most likely to be attributed to insufficiently captured geographical and political factors.

A similar interplay of geographic conditions and cultural factors is found to constitute the extremely low incidence of stunting and wasting in the Lake Victoria region. While fish consumption in combination with an overall food secure situation spurs the growth process of children close to the lake and therefore leads to the much higher body height of children in the Lake Victoria area, the food security situation per se leads to ceteris paribus better wasting rates in the area.

Although these results are already very important for policy making we further examined which single factors contribute most to explaining differences in malnutrition and mortality between Kenyan families. Our analysis reveals that the hiv status of the mother and children's diarrhea status explain the largest part in the variation of stunting outcomes between families while the educational attainment of the mother turns out to be the single most important source in explaining mortality differentials between families.

Our findings demonstrate the relevance of considering and understanding the country specific context, when data on child mortality and malnutrition is analyzed. We do not challenge the epidemiological literature in the sense that we do not question that on the individual level a causal relationship between nutritional and mortality outcomes exists. The analysis raises a serious concern when using children's height status as a reliable proxy for health or income. This is only advisable when geographic, cultural and political contexts are comparable and this is often unlikely to be the case in cross-country or cross-regional analysis.

References

Akin, Johan, David Guilkey, and Barry Popkin, "A child health production function estimated from longitudinal data," *Journal of Development Economics*, 1992, *38*, 323–351.

Alwy, Alwiya and Susanne Schech, "Ethnic Inequalities in Education in Kenya," *International Education Journal*, 2004, *5* (2), 266–274.

Anderson, Carl A., David L. Duffy, Nicholas G. Martin, and Peter M. Visscher, "Estimation of Variance Components for Age at Menarche in Twin Families," *Behavior Genetics*, 2007, *37*, 668–677.

Angeles, Gustavo and Stanislav Kolenikov, "The Use of Discrete Data in Principal Component Analysis: Theory, Simulations, and Application to Socioeconomic Indices," *Proceedings of the American Statistical Association*, 2004.

Becher, Heiko, Olaf Müller, Albrecht Jahn, Adjima Gbangou, Gisela Kynast-Wolf, and Bocar Kouyaté, "Risk factors of infant and child mortality in rural Burkina Faso," *Bulletin of the World Health Organization*, April 2004, *82* (4), 265–273.

Behrman, Jere R. and Anil B. Deolalikar, "Health and Nutrition," in H. Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics No. 1*, Elsevier 1988, pp. 631–711.

Béné, Christophe and Sonja Merten, "Women and Fish-for-Sex: Transactional Sex, HIV/AIDS and Gender in African Fisheries," *World Development*, 2008, *36* (5), 875–899.

Bolstad, William M. and Samuel O. Manda, "Investigating Child Mortality in Malawi Using Family and Community Random Effects: A Bayesian Analysis," *Journal of the American Statistical Association*, March 2001, *96* (453), 12–19.

Bozzoli, Carlos, Angus Deaton, and Climent Quintana-Domeque, "Adult height and childhood disease," *Demography*, forthcoming 2009.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics*, Cambridge University Press, 2005.

Caulfield, Laura E., Stephanie A. Richard, and Robert E. Black, "Undernutrition as an Underlying Cause of Malaria Morbidity and Mortality in Children Less Than Five Years Old," *The American Journal of Tropical Medicine and Hygiene*, 2004, *71* (Suppl 2), 55–63.

CBS, MOH, and ORC Macro, *Final Report on the Kenyan Demographic and Health Survey 2003* 2004.

Chesoni, Atsango, "Ethnic Chauvinism, Xenophobia and Gender Inequality in Kenya," *In: Reading on Inequality in Kenya, vol.1, Society for International Development*, 2006, pp. 195–252.

Cohen, John M., "Ethnicity, Foreign Aid, and Economic Growth in Sub-Saharan Africa: The Case of Kenya," *Harvard - Institute for International Development Working Paper no. 520*, 1995.

Colony and Protectorate of Kenya, *Annual Medical Report ending 31st Decmeber of the year 1922,* 1923.

Cox, David R., "Regression models and life tables," *Journal of the Royal Statistical Society Series B*, 1972, *34*, 187–220.

_ **and David Oakes**, *Analysis of Survival Data*, New York: Chapman and Hall, 1984.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney, "The determinants of Mortality," *Journal of Economic Perspectives*, 2006, *20* (3), 97–120.

Deaton, Angus, "Height, health, and inequality: the distribution of adult height in India," *American Economic Review*, 2008, *98* (2), 468–474.

DiPrete, Thomas A. and Jerry. D. Forristal, "Multilevel Model: Methods and Substance," *Annual Review of Sociology*, 1994, *20*, 331–357.

Fawzi, Wafaie W, M Guillerino Herrera, Donna L Spiegelman, Alawia El Amin, Penelope Nestel, and Kainal A Mohamed, "A prospective study of malnutrition in relation to child mortality in the Sudan," *The American Journal of Clinical Nutrition*, 1997, *65*, 1062–9.

Fearn, H., *An African Economy : A Study of the Economic Development of the Nyanza Province of Kenya, 1903-1953*, Oxford University Press, 1961.

FEWS-Net, "Greater Horn of Africa Food Security Bulletin," Technical Report, FEWS-Net August 2004.

Filmer, Deon and Lant Pritchett, "Estimating Wealth Effects withour Expenditure Data - or Tears: An Application to Educational Enrollments in States of India," *Demography*, 2001, *38* (1), 115–132.

Fogel, Robert, "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy," *American Economic Review*, 1994, *84* (3), 369–395.

Geheb, Kim, Sarah Kallocha, Modesta Medardb, Anne-Therese Nyapendic, Carolyne Lwenyac, and Mercy Kyangwad, "Nile perch and the hungry of Lake Victoria: Gender, status and food in an East African fishery," *Food Policy*, 2008, *33* (1), 85–98.

Goldstein, Harvey, *Multilevel Statistical Models*, Vol. 3, London: Edward Arnold, 2003.

Guo, Guang and German Rodriguez, "Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in Guatemala," *Journal of the American Statistical Association*, 1992, *87* (420), 969–976.

_ **and Hongxin Zhao**, "Multilevel Modeling For Binary Data," *Annual Review of Sociology*, 2000, *26*, 441–62.

Harttgen, Kenneth and Mark Misselhorn, "A Multilevel Approach to Explain Child Mortality and Undernutrition in South Asia and Sub-Saharan Africa," *Ibero-America Institute Working Paper No. 152*, 2006.

Hill, Kenneth, Boaz Cheluget, Siân Curtis, George Bicego, and Mary Mahy, "HIV and Increases in Childhood Mortality in Kenya in the Late 1980s to the Mid-1990s," 2004.

Hox, Joop J., *Applied Multilevel Analysis*, Amsterdam: TT-Publikaties, 1995.

Kenya National Bureau of Statistics, "Basic Report on Well-being in Kenya," Technical Report 2006.

Klasen, Stephan, "Nutrition, Health, and Mortality in Sub Saharan Africa: Is there a Gender Bias?," *Journal of Development Studies*, 1996, *32* (6), 913–932.

_ , "Poverty, undernutrition, and child mortality: Some inter-regional puzzles and their implications for research and policy," *Journal of Economic Inequality*, 2008, *6* (1), 89–115.

MARA, "Mapping Malaria Risk in Africa," Technical Report, MARA 2004.

Marques, Rejane C., Jos Garrofe Drea, Jos V. E. Bernardi, Wanderley R. Bastos, and Olaf Malm, "Maternal fish consumption in the nutrition transition of the Amazon Basin: Growth of exclusively breastfed infants during the first 5 years," *Annals of Human Biology*, 2008, *35* (4), 363–377.

Miguel, Edward, "Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania," *World Politics*, 2004, *56*, 327–362.

_ **and Mary Kay Gugerty**, "Ethnic diversity, social sanctions, and public goods in Kenya," *Journal of Public Economics*, 2005, *89* (11-12), 2325–2368.

Moradi, Alexander, "Towards an Objective Account of Nutrition and Health in Colonial Kenya: A Study of Stature in African Army Recruits and Civilians, 1880-1980," *Journal of Economic History*, forthcoming 2009.

Mosley, W. Henry and Lincoln C. Chen, "An Anyalytical Framework for the Study of Child Survival in Developing Countries," *Population and Development Review*, 1984, *10*, 25–45.

Muhula, Raymond, "Horizontal Inequality and Ethno-Regional Politics in Kenya," *Unpublished Working Paper*, 2008.

Nandy, Shailen, Michelle Irving, David Gordon, S.V. Subramanian, and George Davey Smith, "Poverty, child undernutrition and morbidity: new evidence from India," *Bulletin of the World Health Organization*, March 2005, *83* (3).

Ndugwa, Robert P, Heribert Ramroth, Olaf Müller, Momodou Jasseh, Ali Sié, Bocar Kouyaté, Brian Greenwood, and Heiko Becher, "Comparison of all-cause and malaria-specific mortality from two West African countries with different malaria transmission patterns," *Malaria Journal*, 2008, *71* (15).

Ng'weshemi, Japheth, Mark Urassa, Raphael Isingo, Gabriel Mwaluko, Juliana Ngalula, Ties Boerma, Milly Marston, and Basia Zaba, "HIV Impact on Mother and Child Mortality in Rural Tanzania," *Journal of Acquired Immune Deficiency Syndromes*, 2003, *33*, 393–404.

Nyanjom, Othieno, "Inequality in Kenya's Health Sector," *In: Reading on Inequality in Kenya, vol.1, Society for international development*, 2006, pp. 97–156.

Ochumba, Peter B. O. and David I. Kibaara, "Observations on blue-green algal blooms in the open waters of Lake Victoria," *African Journal of Ecology*, 1989, *27*, 23–34.

Oguttu, Hannington W., Fred W. B. Bugenyi, Heinz Leuenberger, Markus Wolf, and Reinhard Bachofen, "Pollution menacing Lake Victoria : Quantification of point sources around Jinja Town, Uganda," *Water SA*, 2008, *34* (1), 89–98.

Omariba, D. Walter Rasugu, Roderic Beaujot, and Fernando Rajulton, "Determinants of infant and child mortality in Kenya: an analysis controlling for frailty effects," *Population Res. Policy Review*, 2007, *26*, 299–321.

Omumbo, Judith A, Simon I Hay, Carlos A Guerra, and Robert W Snow, "The relationship between the Plasmodium falciparum parasite ratio in childhood and climate estimates of malaria transmission in Kenya," *Malaria Journal*, June 2004, *3*.

Omwega, Reuben N., Richard Abila, and Caroyline Lwenya, "Fishing and poverty levels around Lake Victoria (Kenya)," *Kenya Marine and Fisheries Research Institute Working Paper*, 2003, *10*, 25–45.

Oniang'o, R. K. and A. Komokoti, "Food Habits in Kenya: The Effects of Change and Attendant Methodological Problems," *Appetite*, 1999, *32*, 93–96.

Oster, Emily, "HIV and Sexual Behavior Change: Why not Africa ?," *NBER Working Paper No. 13049*, 2007.

_ , "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa," *NBER Working Paper No. 13610*, 2008.

Pankratz, V. Shane, Mariza de Andrade, and Terry M. Therneau, "Random-Effects Cox Proportional Hazards Model: General Variance Components Methods for Time-to-Event Data," *Genetic Epidemiology*, 2005, *28*, 97–109.

Pelletier, D.L., E.A. Frongillo, D.G. Schroeder, and J.-H-Habicht, "The Effects of Malnutrition on Child Mortality in Developing Countries," *Bulletin of the World Health Organization*, 1995, *73* (4), 443–448.

Peters, Caroline and Rudo Niemeijer, "Protein-Energy Malnutrition and the Home Environment: A study among children in Coast Province, Kenya," *University of Leiden - African Studies Centre Report no. 22/1987*, 1987.

Ramalingaswami, Vulimiri, Urban Jonsson, and Jon Rode, "The Asian enigma," in Peter Adamson, ed., *The Progress of Nations, 1996: The Nations of the World Ranked According to Their Achievements in Child Health, Nutrition, Education, Family Planning, and Progress for Women.*, 1996.

Rosenzweig, Mark R. and Kenneth I. Wolpin, "Heterogeneity, Intrafamily Distribution, and Child Health," *The Journal of Human Resources*, 1988, *13* (4), 437–461.

Sachs, Jeffrey and Pia Malaney, "The Economic and Social Burden of Malaria," *Nature*, 2002, *415*, 680–685.

Sahn, David E. and David Stifel, "Exploring Alternative Measures of Welfare in the Absence of Expenditure Data," *Review of Income and Wealth*, 2003, *49* (4), 463–489.

Sastry, Narayan, "Family-level clustering of childhood mortality risk in Northeast Brazil," *Population Studies*, 1997, *51*, 245–261.

_ , "A nested frailty model for survival data, with an application to the study of child survival in Northeast Brazi," *Journal of the American Statistical Association*, 1997, *92* (438), 426–435.

Scheren, P. A. G. M., H. A. Zanting, and A. M. C. Lemmens, "Estimation of water pollution sources in Lake Victoria, East Africa : Application and elaboration of the rapid assessment methodology," *Journal of environmental management*, 2000, *58* (4), 235–248.

Smith, Lisa C. and Lawrence Haddad, "How Potent Is Economic Growth in Reducing Undernutrition? What Are the Pathways of Impact? New Cross-Country Evidence," *Economic Development and Cultural Change*, October 2002, *51* (1), 55–76.

Snow, Russel W., Eleanor Gouws, J. Omumbo, B. Rapuoda, M. H. Craig, and J. Ouma, "Models to predict the intensity of Plasmodium falciparum transmission: applications to the burden of disease in Kenya," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1998, *92*, 601–606.

Stata, *Stata statistical software: Release 10.0* Stata Corp. 2007.

Steckel, Richard, "Stature and the Standard of Living," *Journal of Economic Literature*, 1995, *3* (4), 1903–1940.

Svedberg, Peter, "Undernutrition in Sub-Saharan Africa: Is there a Gender Bias?," *Journal of Development Studies*, 1990, *26* (3), 469–486.

_ , *Poverty and Undernutrition: Theory, Measurement and Policy*, Oxford: Clarendon Press, 2000.

Therneau, Terry M., *On mixed-effect Cox models, sparse matrices, and modeling data from large pedigrees.* 2006.

Trussel, James and Germán Rodriguez, "Heterogeneity in demographic research," in Julian Adams, David A. Lam, Albert I. Hermanlin, and Peter E. Smouse, eds., *Convergent Issues in Genetics and Demography*, London: Oxford University Press, 1990, pp. 111–132.

UNICEF, "The State of the World's Children," *United Nations Children's Fund*, 2008.

Vaupel, James W., *Kindred lifetimes: frailty models in population genetics*, New York: Oxford University Press,

_ , "Relatives' risks: frailty models of life history data," *Theoretical Population Biology*, 1990, *37* (1), 220–234.

Villamor, Eduardo, Lara Misegades, Maulidi R Fataki, Roger L Mbise, and Wafaie W Fawzi, "Child mortality in relation to HIV infection, nutritional status, and socio-economic background," *International Journal of Epidemiology*, January 2005.

Wainaina, Gituro, "An Inequality Perspective of Education Structure and Performance in Kenya," *In: Reading on Inequality in Kenya, vol.1, Society for international development*, 2006, pp. 157–194.

Wamani, Henry, Anne Nordreghaug, Stefan Peterson, James K. Tumwine, and Thorkild Tyllesk, "Boys are more stunted than girls in Sub-Saharan Africa: A meta-analysis of 16 demographic and health surveys," *BMC Pediatrics*, 2007, *7* (17).

Weinreb, Alexander A., "First Politics, Then Culture: Accounting for Ethnic Differences in Demographic Behaviour in Kenya," *Population and Development Review*, September 2001, *27* (3), 437–467.

Whitworth, Alison and Rob Stephenson, "Birth spacing, sibling rivalry and child mortality in India," *Social Science and Medicine*, 2002, *55*, 2107–2119.

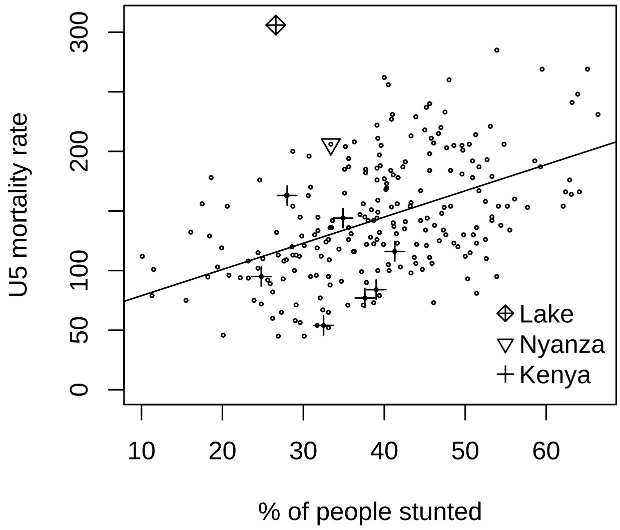
WHO, "WHO Child Growth Standards," *WHO - Backgrounder2*, 2006.

Williamson, Jeffrey G., *Coping with city growth during the British Industrial Revolution*, Cambridge; New York and Melbourne: Cambridge University Press, 1990.

Zaba, Basia, Jimmy Whitworth, Milly Marston, Jessica Nakiyingi, Anthony Ruberantwari, Mark Urassa, Raphaeli Issingo, Gabriel Mwaluko, Sian Floyd, Andrew Nyondo, and Amelia Crampin, "HIV and Mortality of Mothers and Children - Evidence From Cohort Studies in Uganda, Tanzania, and Malawi," *Epidemiology*, 2005, *16*, 275–280.

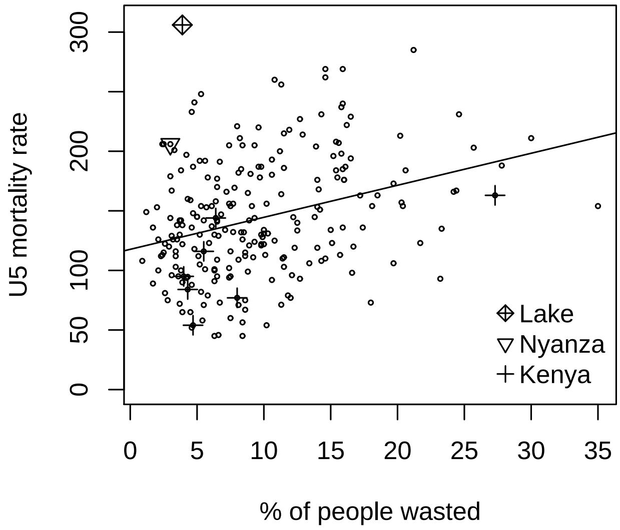
7 Figures and Tables

Figure 1: Stunting rates in SSA



Source: Authors' calculations

Figure 2: Wasting rates in SSA



Source: Authors' calculations

Figure 3: Theoretical Framework

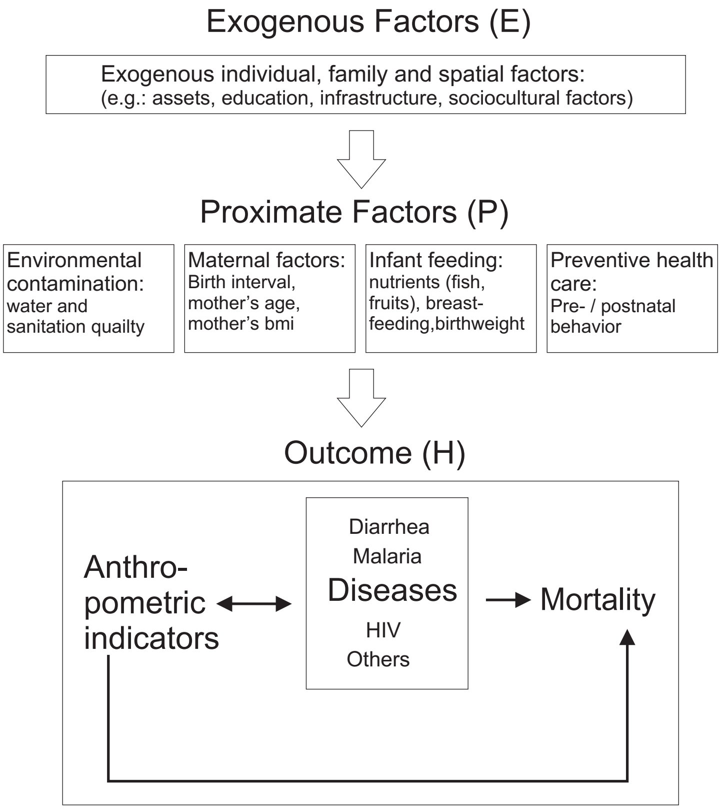
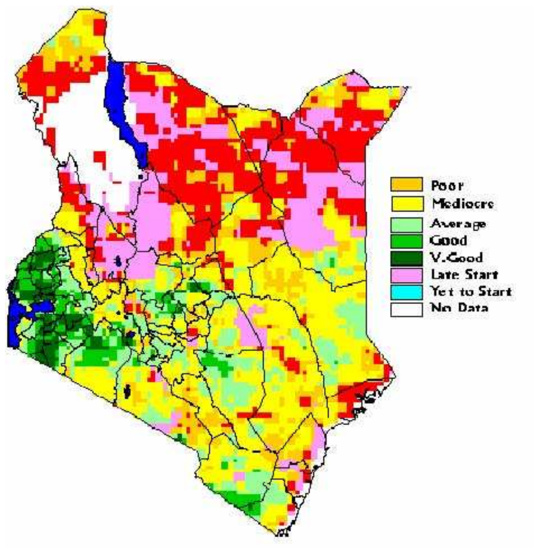
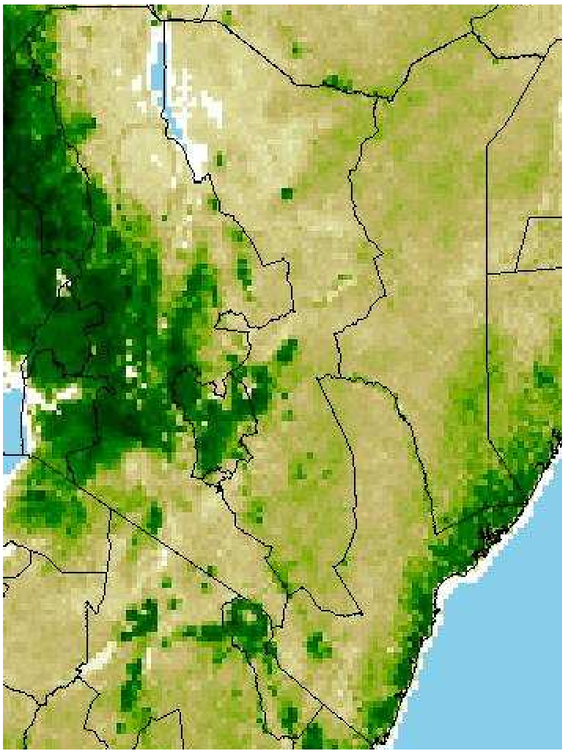


Figure 4: Water Requirement Satisfaction Index (WRSI) Derived Crop Conditions



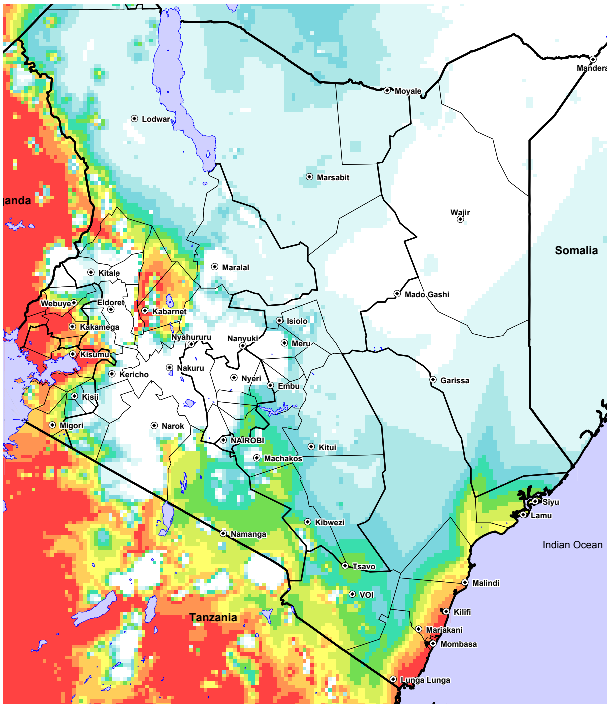
Source: FEWS-Net (2004)

Figure 5: Normalized Difference Vegetation Index - Kenya



Source: FEWS-Net (2004)

Figure 6: Climate Suitability for Endemic Malaria



Source: MARA (2004)

Table 1: Mortality and undernutrition rates in the SSA context

No.	Country	Region	Year	Mortality	Stunting	Wasting	Range	CV
1	Senegal	Tambacounda	2005	200	28.7	11.2	-126	0.31
2	Zambia	Western	2001	201	49.7	3.3	-118	0.22
3	Burkina Faso	Sud-Ouest	2003	203	47.7	25.7	-166	0.22
4	Chad	Zone8	2004	204	35.2	13.9	-122	0.20
5	Cameroon	Nord	2004	205	49.6	8.4	-130	0.27
6	Senegal	Kolda	2005	205	39.6	9.3	-126	0.31
7	Mozambique	Sofala	2003	205	48.6	7.4	-152	0.28
8	**Kenya**	**Nyanza**	**2003**	**206**	**33.4**	**3**	**-152**	**0.40**
9	Mozambique	Tete	2003	206	54.8	2.5	-152	0.28
10	Mozambique	Niassa	2003	206	50.5	2.4	-152	0.28
11	Guinea	Kankan	2005	207	46.1	15.6	-126	0.21
12	Ghana	Upper West	2003	208	36.3	15.4	-133	0.34
13	Guinea	Kindia	2005	211	39.2	8.2	-126	0.21
14	Burkina Faso	Cascades	2003	211	45.8	30	-166	0.22
15	Burkina Faso	Centre-Ouest	2003	213	43.3	20.2	-166	0.22
16	Niger	Tahoua	2006	214	51.3	12.9	-158	0.29
17	Niger	Dosso	2006	215	46.7	11.5	-158	0.29
18	Guinea	NZérékoré	2005	218	45	11.9	-126	0.21
19	Mozambique	Nampula	2003	220	47	9.6	-152	0.28
20	Malawi	Mulanje	2004	221	53.1	8	-109	0.18
21	Mali	Koulikoro	2006	222	39.1	16.2	-154	0.25
22	Mali	Mopti	2006	227	40.9	12.7	-154	0.25
23	Mali	Tombouctou	2006	229	43.9	16.5	-154	0.25
24	Burkina Faso	Nord	2003	231	41	24.6	-166	0.22
25	Niger	Maradi	2006	231	66.4	14.3	-158	0.29
26	Ruanda	East	2005	233	47.5	4.6	-109	0.20
27	Mali	Sikasso	2006	237	45.2	15.8	-154	0.25
28	Chad	Zone5	2004	240	45.6	15.9	-122	0.20
29	Mozambique	Cabo Delgado	2003	241	63.2	4.8	-152	0.28
30	Zambia	Luapula	2001	248	63.9	5.3	-118	0.22
31	Chad	Zone7	2004	256	40.5	11.3	-122	0.20
32	Nigeria	North East	2003	260	48	10.8	-166	0.36
33	Mali	Segou	2006	262	40	14.6	-154	0.25
34	Nigeria	North West	2003	269	59.5	14.6	-166	0.36
35	Niger	Zinder	2006	269	65.1	15.9	-158	0.29
36	Burkina Faso	Sahel	2003	285	53.9	21.2	-166	0.22

Source: Authors' calculations based on latest DHS surveys of respective countries.
Note: CV relates to the coefficient of variation and Range to the difference between the minimum and maximum value within a country. Both measures refer to under-five mortality rates.

Table 2: Classification of Variables Influencing Health and Mortality Outcomes

Variables	**Observed by analysts**	**Unobserved by analysts**
Individual	Health indicators (anthropometric measures, Mortality outcome, reported diarrhea) Health related practices (birth interval, Clinical birth, caesarian section, vaccinations, Retro-viral drugs, duration of breastfeeding, Death of previous child) Age, gender, twin status	Genetic endowment, HIV status, Nutritional intake
Family	Mother's education in years, ethnic belonging, Mother's age at birth, marital status, Mother's BMI and HIV status, Household size, asset possession, Water and sanitation access	Genetic factors, Innate ability for child care, Parental time devoted to child care, General knowledge and mental capability, Income and expenditure levels, Intra-household resource allocation
Community	Geographic location Rural or urban	General disease and health environment Public Infrastructure (availability and quality of education and health care facilities, roads,...), water quality Cultural habits Labor market conditions
District	Malaria suitablity	Political factors
Province	Health access (People per Physician, Health expenditures per capita)	Political factors

Table 3: Description on variables of interest

Variable	Characteristic	Database	Level
Indicators			
Under5-mortality	Child alive (0), died before age of 5 (1)	DHS	Child
Stunting	Stunting z-score[22] (used in regression analysis)	DHS	Child
Wasting	Wasting z-score (used in regression analysis)	DHS	Child
Stunted	Stunting z-score: $>$ -2 standard deviations (sd) (0), $<$ -2 sd (1)	DHS	Child
Wasted	Wasting z-score: $>$ -2 sd (0),$<$ -2 sd (1)	DHS	Child
Child Factors			
Female	Child is male (0), female (1)	DHS	Child
Twin	Child is twin: no (0), yes (1)	DHS	Child
Previous child died	Previously born child of mother died before birth of index child[23]: no (0), yes (1)	DHS	Child
Birth interval	No. of months between birth of index child and birth of preceding child[24]	DHS	Child
Breastfeeding	No. of months child was breastfed	DHS	Child
Diarrhea	Child suffered from diarrhea within last two weeks: no (0), yes (1)	DHS	Child
Family Factors			
Rural	Mother lives in urban area (0), rural area (1)	DHS	Mother
Birth age	Age of mother at birth of index child	DHS	Mother
Married	Single mother (0), married mother (1)	DHS	Mother
HH Size	No. of household members	DHS	Mother
Asset index	Index based on principal component analysis derived by ORC macro[25]	DHS	Mother
Water index	Index: open water (0), uncovered well (1), covered well (2), piped water (3)	DHS	Mother

[22] Z-scores on stunting and wasting are based on the new WHO child growth reference standard (WHO, 2006).

[23] Variable is expected to capture the household specific mortality risk that might arise from potential unobserved genetic, geographic and cultural factors (Sastry, 1997b; Bolstad and Manda, 2001; Whitworth and Stephenson, 2002; Omariba et al., 2007).

[24] To avoid a substantial reduction in the sample size we adopted the same strategy as introduced in Sastry (1997a) and Bolstad and Manda (2001). Preceding birth interval dummies were set to zero for the first born child in a family and in addition a dummy variable was generated to capture the status of a first born child. Likewise, the survival status of the previous child was set to one if the child was of birth order 1.

[25] Since income and expenditure data is not collected in the DHS a common practice is to use available information on the possession of certain household assets to obtain an estimate for the household economic status. ORC Macro International currently relies on Filmer and Pritchett (2001) and applies Principal Component Analysis (PCA) to derive weights for each of the household items taken into account. Many alternative ways to derive weights have been proposed, e.g. using factor analysis as in Sahn and Stifel (2003) or poly-chorical PCA as in Angeles and Kolenikov (2004). For the sake of comparability and reproducibility we decided to stick with the asset index as provided by ORC Macro International.

Toilet index	Index: no facility (0), latrine (1), flush toilet (2)	DHS	Mother
Mother's Education	Years of education of mother	DHS	Mother
Mother's BMI	Mother's BMI status	DHS	Mother
Mother's HIV Status	HIV status of mother based on blood test: HIV negative (0), HIV positive (1)	DHS	Mother
Higher Level Factors			
High diarrhea	Estimated diarrhea prevalence > 50%: no (0), yes (1) (used in mortality regression)	DHS	Community
Malaria Prevalence	Estimated prevalence of malaria	MARA	District
High prevalence	Estimated malaria prevalence > 50%: no (0), yes (1) (used in mortality regression)	MARA	District
People/Physician	Number of people per clinical officer	Nyanjom	Province
Health Expenditures p.c.	Annual national government spending per capita 1995 till 1998	Nyanjom	Province
Pre- Postnatal Behavior			
Clinical birth	Child was born at home (0), born in health center (1)	DHS	Child
Pre birth visits	No. of times mother visited health center prior to birth of index child	DHS	Child
Caesarian	Child was born without section (0), under caesarian section (1)	DHS	Child
Retro viral drugs	Drug intake to prevent HIV transmission to child during pregnancy: no (0), yes (1)	DHS	Child
Vaccination index	Newborn (0-1 months) received no vaccination (1), at least one vaccination (1)	DHS	Child
	Infant (2-11 months) received no vaccination (0), at least one vaccination (1)	DHS	Child
	Child (12-59 months) did not receive all vaccinations (0), all vaccinations (1)	DHS	Child
Nutritional Indicators			
Protein	Child receives meat, fish or eggs at least once a week: no (0), yes (1)	DHS	Child
Grain	Child receives food made from local grain at least once a week: no (0), yes (1)	DHS	Child
Vitamin A	Child receives vitamin A rich fruits at least once a week: no (0), yes (1)	DHS	Child

Source: Based on KDHS 2003

Table 4: Descriptive statistics on variables of interest

	Lake Region	Rest of Kenya	Lake Region		Kenya	Nairobi	Central	Coast	East.	Nyanza	Rift Val.	West.	North East.
			Luo	Other Ethn.									
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**	**XI**	**XII**	**XIII**
Indicators													
Under5-mortality	30.9%	9.6%	32.6%	24.9%	11.5%	8.4%	4.9%	10.3%	9.6%	22.4%	7.3%	16.0%	12.1%
Stunted	26.7%	37.5%	24.2%	34.6%	36.8%	24.8%	35.2%	42.0%	39.8%	34.6%	39.0%	36.6%	26.2%
Wasted	4.9%	6.1%	4.2%	6.9%	6.0%	4.3%	4.6%	5.2%	3.9%	2.8%	7.7%	7.5%	27.2%
Child Factors													
Female	48.5%	48.6%	47.2%	54.2%	48.6%	49.8%	48.8%	47.4%	48.2%	48.0%	48.9%	50.3%	45.8%
Twin	3.0%	4.4%	2.7%	4.3%	4.2%	3.5%	4.1%	6.1%	3.7%	4.8%	3.2%	5.0%	5.5%
Prev. child died	27.2%	14.8%	29.8%	15.0%	16.2%	15.3%	7.2%	17.9%	11.6%	26.4%	12.2%	17.0%	21.1%
Birth interval	31.23	36.78	29.70	38.51	36.15	42.81	46.93	36.50	39.32	32.01	35.57	33.00	27.63
Breastfeeding	13.62	14.54	13.42	14.32	14.46	13.12	15.23	14.36	15.85	13.97	14.23	14.61	11.50
Diarrhea	21.7%	15.7%	20.8%	24.7%	16.1%	13.3%	6.8%	22.0%	12.6%	17.3%	17.0%	23.8%	13.6%
Family Factors													
Rural	91.2%	81.6%	92.3%	86.4%	82.6%	0.0%	87.1%	68.3%	95.9%	92.7%	84.2%	92.7%	82.5%
Birth age	25.14	25.36	24.96	25.90	25.33	24.53	25.68	24.76	26.14	24.87	25.50	25.10	26.12
Married	72.0%	77.7%	69.4%	83.5%	77.1%	73.0%	75.5%	75.8%	81.4%	70.2%	75.5%	86.2%	90.2%
HH Size	5.56	5.99	5.55	5.63	5.94	5.04	5.28	6.63	6.48	5.86	5.96	5.74	6.33
Asset index	-0.59	-0.34	-0.62	-0.45	-0.37	1.55	-0.13	-0.31	-0.54	-0.62	-0.41	-0.68	-0.87
Water index	1.89	2.04	1.90	1.84	2.02	3.81	2.20	2.64	1.92	1.67	1.96	1.60	1.49
Toilet index	1.50	1.86	1.41	1.90	1.82	2.60	2.07	1.59	1.88	1.67	1.71	1.99	1.11
Mother's Education	5.76	5.98	5.68	6.12	5.96	8.86	8.12	3.96	6.35	5.97	5.65	6.04	0.18
Mother's BMI	22.06	22.44	21.95	22.54	22.40	25.07	23.45	22.51	22.16	22.18	22.10	22.12	20.00
Mother's HIV Stat.	23.7%	8.8%	26.3%	10.6%	10.5%	18.5%	7.8%	7.2%	9.8%	18.4%	8.7%	5.6%	0.0%
Higher Level Factors													
Malaria Prevalence	48.5%	23.0%	49.5%	43.4%	25.7%	4.7%	6.8%	27.9%	22.2%	48.4%	18.3%	35.6%	7.6%

	Lake Region	**Rest of Kenya**	**Lake Region**		**Kenya**	**Nairobi**	**Central**	**Coast**	**East.**	**Nyanza**	**Rift Val.**	**West.**	**North East.**
			Luo	**Other Ethn.**									
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**	**XI**	**XII**	**XIII**
People/Physician	1102	941	1100	1110	958	813	705	819	848	1100	917	1114	1646
Health Exp. p.c.	1.05	1.61	1.03	1.13	1.55	1.37	1.77	2.43	1.50	1.03	1.70	1.17	2.50
Pre- Postnatal Behavior													
Clinical birth	32.8%	39.7%	31.1%	38.9%	39.2%	76.6%	64.2%	29.8%	36.6%	35.3%	35.4%	28.3%	7.9%
Pre birth visits	4.18	6.40	3.33	6.77	6.23	7.03	6.47	7.48	5.52	6.34	6.75	5.32	1.17
Caesarian	2.5%	4.1%	1.7%	5.0%	3.9%	10.7%	6.2%	3.8%	4.3%	2.0%	3.4%	2.2%	2.7%
Retro viral drugs	38.7%	36.2%	38.9%	37.8%	36.5%	51.1%	44.2%	25.4%	26.3%	39.8%	36.1%	41.6%	1.9%
Vaccination index	56.3%	70.8%	49.8%	76.2%	69.7%	73.9%	86.5%	74.5%	77.4%	56.8%	71.3%	62.9%	20.0%
Nutritional Indicators													
Protein	35.9%	22.0%	38.4%	26.8%	23.0%	34.1%	14.9%	29.7%	23.9%	31.1%	16.8%	25.0%	9.1%
Grain	76.2%	67.0%	76.6%	74.6%	67.6%	66.8%	75.1%	59.2%	80.1%	75.5%	55.8%	72.7%	41.7%
Vitamin A	29.5%	22.3%	30.0%	27.7%	22.8%	30.1%	16.8%	24.8%	33.8%	28.7%	16.7%	18.2%	6.6%

Source: Authors' calculations based on KDHS 2003

Table 5: Determinants of stunting

	Model I				Model II				Model III			
	Est.	**SE**	**p-value**		**Est.**	**SE**	**p-value**		**Est.**	**SE**	**p-value**	
Fixed Effects												
Female	0.315	0.078	0.000	***	0.296	0.071	0.000	***	0.295	0.071	0.000	***
Twins	-0.760	0.240	0.002	***	-0.605	0.255	0.018	**	-0.614	0.254	0.016	**
Rural	-0.002	0.130	0.986		0.003	0.142	0.981		-0.008	0.159	0.962	
log(Mother's BMI)	0.754	0.256	0.003	***	0.769	0.281	0.006	***	0.711	0.281	0.011	***
Asset index	0.197	0.089	0.026	**	0.192	0.095	0.045	**	0.182	0.101	0.073	*
Mother's education	0.111	0.037	0.003	***	0.096	0.041	0.020	**	0.103	0.042	0.013	***
Water index	0.119	0.037	0.001	***	0.108	0.041	0.008	***	0.102	0.043	0.017	**
Toilet index	0.006	0.110	0.956		0.017	0.121	0.886		0.021	0.127	0.870	
Birth age	0.010	0.012	0.416		0.011	0.013	0.391		0.010	0.013	0.463	
Birth order	-0.015	0.035	0.663		-0.037	0.038	0.329		-0.031	0.038	0.413	
First child	0.221	0.114	0.053	**	0.145	0.104	0.163		0.150	0.104	0.148	
Birth interval	0.003	0.002	0.100	*	0.003	0.002	0.191		0.003	0.002	0.134	
Previousdead	0.080	0.200	0.689		0.139	0.192	0.469		0.161	0.191	0.397	
Luo	-0.312	0.167	0.061	*	-0.288	0.182	0.114		-0.244	0.188	0.195	
HIV	0.156	0.158	0.324		0.140	0.170	0.411		0.142	0.169	0.398	
High prevalence	0.127	0.164	0.440		0.066	0.185	0.720		0.050	0.208	0.810	
Diarrhea	-0.070	0.106	0.508		-0.068	0.099	0.495		-0.063	0.099	0.527	
log(People/Physician)	1.298	0.225	0.000	***	1.115	0.249	0.000	***	1.113	0.278	0.000	***
Health Expenditures p.c.	0.157	0.113	0.166		0.121	0.124	0.330		-0.020	0.228	0.931	
Lake Region	0.459	0.214	0.032	**	0.487	0.234	0.037	**	0.473	0.252	0.061	*
Constant	-13.748	1.882	0.000	***	-12.370	2.073	0.000	***	-12.051	2.266	0.000	***
Random Effects												
Cluster random effect									0.21			
Family random effect					1.21				1.01			
Individual random effect					1.30				1.30			
Observations	1704				1704				1704			
Integrated log Likelihood	-3199				-3158				-3152			
LR-Test chisq(1)					153			***	13			***
LR-Test chisq(2)									660			***

Source: Authors' calculations based on KDHS 2003

Table 6: Determinants of wasting

	Model I				Model II				Model III			
	Est.	**SE**	**p-value**		**Est.**	**SE**	**p-value**		**Est.**	**SE**	**p-value**	
Fixed Effects												
Female	0.001	0.061	0.987		-0.003	0.057	0.954		0.000	0.057	0.997	
Twins	-0.086	0.188	0.648		-0.096	0.202	0.635		-0.086	0.202	0.671	
Rural	-0.058	0.102	0.570		-0.050	0.112	0.656		-0.049	0.119	0.681	
log(Mother's BMI)	13.343	4.202	0.002	***	12.364	4.747	0.009	***	11.819	4.743	0.013	**
log(Mother's BMI 2)	-1.935	0.668	0.004	***	-1.793	0.753	0.017	**	-1.709	0.753	0.023	**
Asset index	-0.057	0.070	0.411		-0.037	0.075	0.625		-0.034	0.078	0.660	
Mother's education	0.039	0.030	0.183		0.040	0.033	0.222		0.037	0.033	0.259	
Water index	0.040	0.029	0.167		0.050	0.032	0.115		0.049	0.033	0.137	
Toilet index	0.176	0.087	0.043	**	0.173	0.096	0.071	*	0.171	0.098	0.083	*
Birth age	0.009	0.010	0.370		0.010	0.010	0.340		0.010	0.010	0.327	
Birth order	-0.014	0.028	0.619		-0.013	0.030	0.663		-0.016	0.030	0.591	
First child	0.116	0.090	0.196		0.072	0.084	0.390		0.066	0.084	0.427	
Birth interval	0.001	0.002	0.487		0.000	0.002	0.753		0.000	0.002	0.838	
Previousdead	0.412	0.357	0.459		0.331	0.453	0.451		0.332	0.454	0.451	
Luo	0.136	0.130	0.298		0.104	0.142	0.467		0.086	0.145	0.554	
HIV	-0.123	0.124	0.322		-0.102	0.134	0.444		-0.112	0.133	0.402	
High prevalence	0.205	0.129	0.113		0.186	0.145	0.199		0.196	0.154	0.201	
Diarrhea	-0.128	0.083	0.124		-0.077	0.079	0.335		-0.074	0.079	0.351	
log(People/Physician)	-0.848	0.177	0.000	***	-0.808	0.196	0.000	***	-0.820	0.207	0.000	***
Health Expenditures p.c.	0.014	0.143	0.919		0.003	0.158	0.983		-0.014	0.168	0.933	
Lake Region	0.155	0.167	0.356		0.167	0.183	0.363		0.184	0.190	0.335	
Distance to Lake Region	-0.001	0.000	0.084	*	-0.001	0.000	0.080	*	-0.001	0.000	0.103	*
Constant	-17.531	6.735	0.009	***	-16.123	7.622	0.034	**	-15.12	7.639	0.048	**
Random Effects												
Cluster random effect									0.06			
Family random effect					0.67				0.61			
Individual random effect					0.89				0.89			
Observations	1701				1701				1701			
Integrated log Likelihood	-2775				-2773				-2771			
LR-Test chisq(1)					97.6			***	3.5			***
LR-Test chisq(2)									101.1			***

Source: Authors' calculations based on KDHS 2003

Table 7: Determinants of under5-mortality

	Model I				Model II				Model III			
	Hazard	**SE**	**p-value**		**Hazard**	**SE**	**p-value**		**Hazard**	**SE**	**p-value**	
Fixed Effects												
Female	0.84	0.08	0.04	**	0.82	0.09	0.03	**	0.82	0.09	0.03	**
Twins	1.51	0.19	0.03	**	1.57	0.22	0.04	**	1.53	0.22	0.05	**
Rural	1.09	0.15	0.57		1.09	0.18	0.63		1.14	0.19	0.50	
log(Mother's BMI)	2.35	4.31	0.84		0.37	5.82	0.87		0.23	5.71	0.80	
log(Mother's BMI 2)	1.08	0.67	0.91		1.53	0.91	0.64		1.64	0.89	0.58	
Asset index	0.87	0.11	0.19		0.86	0.12	0.23		0.89	0.13	0.37	
Mother's education	0.55	0.11	0.00	***	0.49	0.13	0.00	***	0.47	0.14	0.00	***
Mother's education 2	1.07	0.03	0.01	***	1.09	0.03	0.01	***	1.10	0.03	0.00	***
Water index	1.09	0.04	0.04	**	1.08	0.05	0.13		1.08	0.05	0.12	
Toilet index	0.99	0.12	0.93		1.01	0.14	0.97		0.98	0.14	0.86	
Birth age	0.88	0.05	0.01	***	0.85	0.06	0.00	***	0.83	0.06	0.00	***
Birth age 2	1.00	0.00	0.16		1.00	0.00	0.10	*	1.00	0.00	0.04	**
Birth order	1.05	0.03	0.18		1.00	0.04	0.95		1.10	0.10	0.31	
First child	0.73	0.13	0.02	**	0.67	0.14	0.01	***	0.73	0.17	0.06	*
Birth interval	0.96	0.01	0.00	***	0.96	0.01	0.00	***	0.96	0.01	0.00	***
Birth interval2	1.00	0.00	0.00	***	1.00	0.00	0.00	***	1.00	0.00	0.00	***
Previousdead	1.43	0.11	0.00	***	1.12	0.13	0.38		1.14	0.13	0.30	
Luo	1.54	0.14	0.00	***	1.59	0.17	0.01	***	1.57	0.18	0.01	***
HIV	1.51	0.13	0.00	***	1.61	0.16	0.00	***	1.54	0.16	0.01	***
High prevalence	1.35	0.13	0.02	**	1.37	0.16	0.06	*	1.36	0.19	0.10	*
High Diarrhea	1.94	0.21	0.00	***	1.95	0.26	0.01	***	1.94	0.28	0.02	**
log(People/Physician)	1.95	0.23	0.00	***	2.13	0.27	0.01	***	2.13	0.30	0.01	***
Health Expenditures p.c.	0.74	0.12	0.01	***	0.75	0.14	0.04	**	0.74	0.15	0.04	**
Lake Region	1.28	0.15	0.09	*	1.43	0.19	0.06	*	1.43	0.21	0.09	*
Random Effects												
Cluster random effect									0.16			
Family random effect					0.60				0.42			
Observations	2697				2697				2697			
Integrated log Likelihood	-4293				-4281				-4277			
LR-Test chisq(1)					12.57			***	3.85			**
LR-Test chisq(2)									16.42			***

Source: Authors' calculations based on KDHS 2003

Table 8: Variance of family random effect

	Stunting regression		**Wasting regression**		**Under-5 mortality regression**	
	Covariate included in null model	Covariate removed from full model	Covariate included in null model	Covariate removed from full model	Covariate included in null model	Covariate removed from full model
None	**1.657**	**1.213**	**1.102**	**0.673**	**0.903**	**0.596**
Female	1.640	1.233	1.102	0.673	0.903	0.582
Twins	1.650	1.222	1.102	0.673	0.893	0.605
Rural	1.611	1.211	1.089	0.673	0.887	0.596
log(Mother's BMI)	1.614	1.192	1.013	0.715	0.912	0.598
Asset index	1.581	1.218	1.054	0.673	0.846	0.602
Mother's education	1.618	1.224	1.031	0.673	0.758	0.650
Water index	1.615	1.216	1.082	0.676	0.895	0.617
Toilet index	1.638	1.208	1.035	0.675	0.846	0.597
Birth age	1.656	1.212	1.102	0.673	1.448	0.608
Birth order	1.646	1.210	1.096	0.673	1.218	0.602
First child	1.649	1.220	1.099	0.677	0.902	0.572
Birth interval	1.645	1.229	1.092	0.679	0.907	0.682
Previousdead	1.655	1.224	1.102	0.681	0.650	0.650
Luo	1.656	1.216	1.082	0.676	0.706	0.629
HIV	1.589	1.322	0.927	0.792	0.870	0.611
High prevalence	1.657	1.212	1.101	0.675	0.856	0.611
High Diarrhea	1.551	1.385	0.982	0.745	0.901	0.616
log(People/Physic.)	1.637	1.260	1.054	0.694	0.798	0.596
Health Expend. p.c.	1.658	1.214	1.062	0.672	0.867	0.623
Lake Region	1.653	1.217	1.061	0.674	0.830	0.590

Source: Authors' calculations based on KDHS 2003