Available online at www.sciencedirect.com

SCIENCE DIRECT

Food Policy 30 (2005) 584–611

FOOD POLICY

www.elsevier.com/locate/foodpol

Child hunger in the developing world: An analysis of environmental and social correlates

Deborah Balk [a,*], Adam Storeygard [a,1], Marc Levy [a,1], Joanne Gaskell [b,2], Manohar Sharma [c,3], Rafael Flor [d,4]

[a] Center for International Earth Science Information Network, Columbia University, P.O. Box 1000, Palisades, NY 10964, United States

[b] Interdisciplinary Program in Environment and Resources, 397 Panama Mall, Mitchell Bldg Rm. 132, Stanford University, Stanford, CA 94305-2210, United States

[c] International Food Policy Research Institute, 2033 K Street NW, Washington, DC 20006-1002, United States

[d] Tropical Agriculture Program, Earth Institute at Columbia University, P.O. Box 1000, 61 Route 9W, Lamont Hall 2E, Palisades, NY 10964, United States

Abstract

Using two complementary methods in a framework that allows incorporating both environmental and household-level factors, we explore the correlates of underweight status among children. We use individual children as the units of analysis in 19 African countries, and subnational survey strata in 43 African, Asian and Latin American countries. We consider the relationship between household-level demographic and health survey data, environmental factors from external geospatial data sets and two indicators of malnutrition among children aged 1–3, deviations from the international standards of weight-for-age and height-for-age. We discuss methods for data integration. In general, household determinants explain more variation than environmental factors, perhaps partly due to more error-prone measurement at the community level. Among individual children, some measures of agricultural capacity are related to lower incidence of child hunger, while among regions, mea-

* Corresponding author. Tel.: +1 845 365 8988; fax: +1 845 365 8922.

E-mail addresses: dbalk@ciesin.columbia.edu (D. Balk), astoreyg@ciesin.columbia.edu (A. Storeygard), mlevy@ciesin.columbia.edu (M. Levy), jgaskell@stanford.edu (J. Gaskell), m.sharma@cgiar.org (M. Sharma), rflor@iri.columbia.edu (R. Flor).

1 Tel.: +1 845 365 8988; fax: +1 845 365 8922.

2 Tel.: +1 650 725 5413.

3 Tel.: +1 202 862 5600; fax: +1 202 467 4439.

4 Tel.: +1 845 680 4531; fax: +1 845 680 4870.

doi:10.1016/j.foodpol.2005.10.007

sures relating to urbanness and population density show a stronger relationship. We give recommendations for further study, data collection and policy making.

Keywords: Child hunger; Underweight; Household; Environmental factors; Spatial analysis

Introduction

Despite the importance of understanding and reducing child hunger, the study of childhood malnutrition has been limited to a degree by division into disciplinary approaches to analysis. Social scientists have shown that individual and household-level factors such as birth order or parental schooling are influential (cf. Charmarbagwala et al., 2004). Nutritionists have shown that there are hereditary and gender differences in the uptake of particular nutrients and ultimately child nutritional outcomes (cf. Payne, 1990). Soil and climate scientists have demonstrated the strong relationship between climate and soil potential for agricultural sustainability and food production (Sanchez, 2002). While conceptual frameworks link these factors closely together (e.g., the UNICEF framework adapted from Smith and Haddad, 2000), few studies systematically assess them in combination. Here, we attempt to close that gap.

Addressing environmental and socio-economic and demographic determinants of child hunger in an integrated fashion requires the adoption of a spatial framework of data integration. Environmental data tend to be spatially referenced but not included in studies of household-level survey data. Thus, this analysis (1) converts spatial information into survey units and (2) converts survey information into different spatial units. This method also presents us with an ability to improve the measurement of common, implicitly spatial factors by treating them in explicitly spatial terms, for example, urban areas, disease vectors and distances to important access points.

The United Nations Millennium Project Hunger Task Force (HTF) recently recommended complementary feeding along with continued breast-feeding until age 2 as one of three pillars to reduce malnutrition in children aged under 5. Similarly, it recommended reducing vulnerability through productive safety nets (e.g., local early warning systems), increasing income and access to markets for the poor (e.g., increase capacity of rural towns and market-related infrastructure, and increase non-farm-dependent income generation) and restoring and conserving essential natural resources for food security (Sanchez et al., 2005). This study allows for a preliminary analysis of how the current state of some of these factors affect child hunger.

Study design and data

Anthropometric indicators from household survey data have long been used as measures of health and nutritional status (Tanner, 1981). Anthropometric indicators are relatively easy to determine for every child in large-scale representative surveys. The data can be compared across countries, "There is evidence that the growth in height and weight of well fed, healthy children, or children who experience unconstrained growth, from different ethnic backgrounds and different continents is reasonably similar, at least up to 5 years of age. While it is accepted that there are variations in the growth patterns of children from different racial or ethnic groups in developed countries, these are relatively minor compared with the large worldwide variation that relates to health, nutrition, and socio-economic status" (WHO, 1995).[5]

Individual-level anthropometric data collected through surveys may also be combined in analysis with all the other data found in those surveys. Several conceptual frameworks have been proposed, placing causes of child malnutrition in a theoretical context (UNICEF, 1990, 1998, further adapted by Engle et al., 1999 and Smith and Haddad, 2000). The Smith and Haddad framework (p. 6), for example, encompasses socio-economic and biological factors operating at household, community and institutional levels. In the past, however, household surveys used for testing this framework have not contained (or been linked to) any information on many of its components.

In the past 15 years, Demographic and Health Surveys (DHS) conducted in over 30 countries have begun to include information about the geographic location of survey respondents. This can then be coupled with factors that are geographically determined or recorded, such as soil quality, agricultural production, population density and distance to a port, to be analysed in combination with the survey data (cf. Balk et al., 2004). These variables are thought to exert influence over and above those effects which are mediated by household characteristics (such as parental schooling or wealth) but in the absence of a geospatial framework are generally explained as regional affects, or simply unexplained variance. Thus, we consider the relationship between child malnutrition, as measured by weight-for-age and height-for-age Z scores (WHO, 1995), and selected biophysical and macro-demographic variables joined to the anthropometric data via geographic location.

We conducted the analysis in two phases, corresponding to two levels of analysis, both of which have inherent strengths and weaknesses. Individual-level analysis permits a much fuller examination of behavioural factors than could be considered at a subnational level. The number of observations is great ($n = 33{,}000$ children), allowing for practically unlimited permutations of variables in form or in combination with others (e.g., interactions). However, it is geographically limited to parts of Africa. Additionally, policy is generally targeted at a level above the household, and while individual-level analyses may be a critical piece of constructing policies, they may also fall short of understanding aspects of the problem that can be best summarized at a regional scale.

The meso-scale analysis has greater geographic breadth and provides a valuable tool for analysts designing policies for whole provinces. But it poses additional pitfalls. There are far fewer units in the analysis ($n = 319$ subnational units), so the risk of reducing the power of the model by introducing too many variables (such as a full suite of interaction terms) is increased. Correspondingly, aggregation removes the variance within each subnational region both among the variables native to the household survey, which must be summarized, and among the variables external to the survey, which must be recalculated

[5] Several studies have concluded that there are no biological differences in ethnicity that might bias comparison of children in a global analysis of weight-for-age status (Habicht et al., 1974; Graitcher and Gentry, 1981; Whitehead and Paul, 1984).

for the region as a whole. Thus, the ecological fallacy must be accounted for in interpreting any results (Freedman, 2001).[6]

By conducting micro- and meso-scale analyses of hunger, we can both benefit from behavioural insight and address policy concerns at a regional scale.

Phase I. Micro-level analysis of determinants of underweight children in Africa

In Phase I, we integrated household-survey data with selected geospatial variables to allow for a child-level analysis of underweight status.[7] The household data come from 19 DHS conducted in African countries in the period 1990–2002. Where multiple surveys were carried out over this time period in a given country, we weighed considerations of data comparability, currency and sample size in selecting one. The DHS are nationally representative.[8] The sample design is a probabilistic two-stage sample where enumeration areas (EAs) are randomly selected with probability proportional to their size. Households within the selected EAs are randomly selected with equal probability, and sampling weights are assigned to individuals. The DHS sampling manual (Macro International Inc., 1996) presents a thorough review of sampling methodology.

In each of these countries, geographic latitude and longitude coordinates were recorded for each survey cluster. A cluster typically corresponds to 20–30 sampled households in a single EA. In all but three surveys, the approximate population centroid of each cluster was determined using a handheld Global Positioning System (GPS) device with an expected accuracy of 100 m or better. In Central African Republic, Chad and Nigeria, clusters were georeferenced using information from published gazetteers, which tend to have lower accuracy. Point locations are also affected by precision problems more significant than these accuracy issues. Households sampled within a cluster can be spread throughout an EA, although the clustered sampling methods used in many countries minimize this. The problem is likely to be largest in sparsely populated rural areas.

For the purposes of this study, we recorded several variables for each cluster location. The integrated child-level data set included all those African surveys listed in Table 1 as having geographic data. The survey clusters are shown in Fig. 1, with prevalence of hunger at the subnational level.[9] We derived the spatial variables from numerous data sets (Table 2). We integrated the GPS-point locations of the sampling clusters with various spatial data (in points, polygons, lines or grid format) by association with the nearest applicable feature (e.g., distance to nearest port), or averaging over features within a given distance

[6] An alternative approach would have been to use a multilevel model; however, this would have entailed removing over half of the regions of interest. Further analysis could pursue this approach to determine whether the tests conducted here would be equally robust with this method for the subset of the study region, assuming that the omitted regions were not systematic in their selection out of full-information sample.

[7] A child is considered underweight (low weight) when its weight is below minus 2 SD from the U.S. National Centre for Health Statistics/WHO international reference median value for its age.

[8] The Mali and Niger surveys exclude remote populations, totalling 2.6% and 4.7% of their populations, respectively. The Kenya survey excludes North-eastern Province and four districts in the northern portions of Eastern and Rift Valley Provinces, totalling less than 4% of the national population. Residents of refugee camps were not surveyed in Guinea (this is likely true in other applicable countries, such as Tanzania, but not explicitly noted).

[9] Cluster-level indices of underweight prevalence were not constructed because they are not supported by the sampling frame. See related discussion in Balk et al. (2004).

Table 1
Countries and data sets studied

Region Country	Year of survey	Number of children aged 1–3 with valid anthropometry data collected
Africa		
Benin[a]	2001	1556
Burkina Faso[a]	1998–1999	1418
Cameroon[a]	1998	1072
Central African Republic[a]	1994–1995	1377
Chad[a]	1996–1997	2022
Côte d'Ivoire[a]	1994	2027
Ethiopia[a]	2000	3334
Ghana[a]	1998	1062
Guinea[a]	1999	1173
Kenya[a]	1998	1886
Madagascar[a]	1997	1786
Malawi[a]	2000	3795
Mali[a]	2001	3607
Namibia[a]	2000	1153
Niger[a]	1998	2357
Nigeria[a]	1990	2125
Tanzania[a]	1999	968
Togo[a]	1998	2071
Zimbabwe[a]	1999	1073
Comoros	1996	574
Egypt	2000	4046
Eritrea	1995	1268
Gabon	2000	1358
Morocco	1992	1783
Mozambique	1997	1792
Rwanda	2000	2228
Uganda	1995	2230
Zambia	1996	2323
Asia		
Bangladesh[a]	1999–2000	1997
Cambodia[a]	2000	1265
Nepal[a]	2001	2342
India	1998–1999	15277
Kazakhstan	1999	223
Kyrgyz Republic	1997	615
Pakistan	1990–1991	1636
Uzbekistan	1996	647
Latin American and the Caribbean		
Haiti[a]	2000	2263
Peru[a]	2000	4502
Bolivia	1998	2362
Brazil	1996	1612
Colombia	2000	1681
Dominican Republic	2002	3811
Guatemala	1995	3455
Nicaragua	1997–1998	2607

[a] Geographic data were available for surveys.

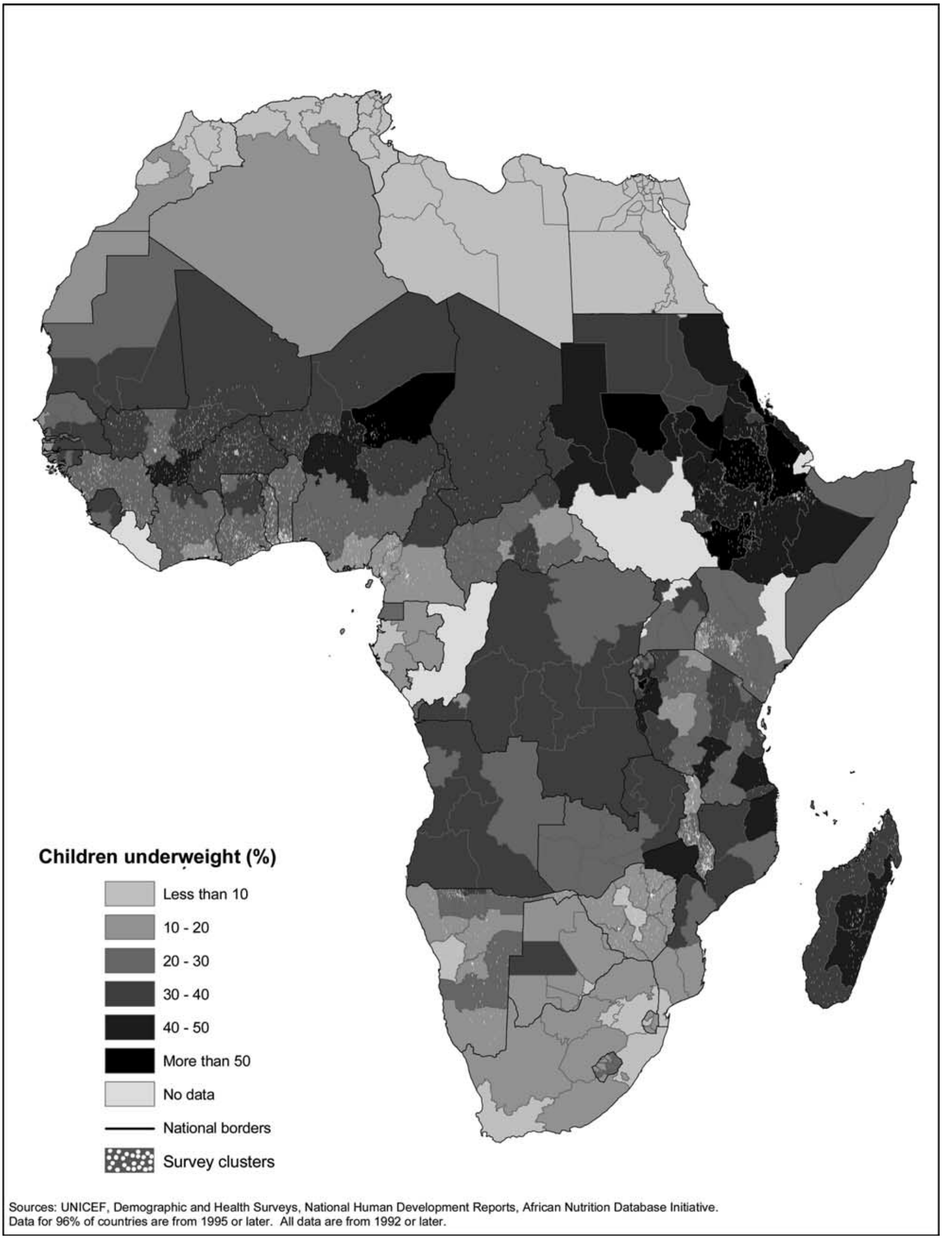

Fig. 1. Prevalence of child malnutrition, with survey clusters.

(e.g., average population within a 10-km radius of the cluster point location), producing an integrated data set. We used averaging over a given distance especially in cases of geographic data with high variability over short distances (e.g., population) both to account for the precision problems described above and to spread out the area over which an effect

Table 2
Variables derived from environmental data sets that were linked with the survey data

Variable	Source	Description[a]	Association with cluster points[a]	Resolution (min[b])
Population				
Population density	CIESIN and CIAT (2003)	In persons per km^2 of land area	Averaged within a 10-km radius of the cluster point	2.5
Urban proximity	Based on CIESIN et al. (2004) urban areas data	Modelled travel time in hours to nearest town of population >25k, >50k, >100k and >500k without crossing water or international borders	Averaged within (A) 10 km and (B) 25 km	0.5
Conflict	CIESIN, unpublished data, based on Gleditsch et al. (2002)	Number of calendar years with political conflict of high, medium and low intensity, since 1975 and 1990	Value at cluster point	30
Infrastructure				
Seaport proximity	Based on World Bank data	Distance to nearest port of any size (km)	Value at centre of quarter-degree grid cell containing cluster point	15
Road density	Based on NIMA (1995)	In km of roads per km^2 of land area	Within 100-km radius	n/a
Climate and cropping				
Cropping intensity	Ramankutty and Foley (1998)	Percentage of 5-min cell cropped	Averaged within (A) 10-km, (B) 25-km and (C) 50-km radius	5
Pasture intensity	Unpublished data from Foley et al. (2003)	Percentage of 5-min cell in pasture	Averaged within (A) 10-km, (B) 25-km and (C) 50-km radius	5
Tree cover density	DeFries et al. (2000)	Percentage of half-min cell with tree cover	Averaged within (A) 10-km, (B) 25-km and (C) 50-km radius	5
Cereal production	IFPRI, unpublished data, based on Ramankutty and Foley (1998) and FAO (2003)	In calories per capita	(A) value at cluster point, (B) averaged within 25-km radius	5
Average rainfall	IIASA-FAO (2000)	In mm	Value at cluster point	30
Length of growing period	IIASA-FAO (2000)	In months	Value at cluster point	30
Farming system[a]	Dixon et al. (2001)	Classes	Value at cluster point	n/a
Soil constraints	IFPRI, unpublished data, based on Sanchez et al. (2005) and FAO (1995)	Percentage of 5-min grid cell with low CEC (high leaching potential), aluminium toxicity, low potassium reserves, shallow or gravelly, or organic soils	Values at cluster point (one for each constraint)	5
Physiography				
Elevation	Rabus et al. (2003)	In m above sea level	Value at cluster point	0.5
Slope	Rabus et al. (2003)	In %	Value at cluster point	0.5

Table 2 (*continued*)

Variable	Source	Description[a]	Association with cluster points[a]	Resolution (min[b])
Roughness[a]	Nelson, unpublished data, 2001	Classes based on Meybeck et al. (2001)	Value at cluster point	5
Disease ecology				
Index of stability of malaria transmission	Kiszewski et al. (2004)	Composite of epidemiological and environmental data	Value at cluster point	30

[a] Some variants were removed during preliminary variable selection in favour of other variables listed.
[b] 1 min is approximately equal to 1.85 km at the equator.

might occur hypothetically. For example, spatially discrete phenomena, such as cropland, are likely to spill over in important ways (e.g., through economic activity and disease transmission) into nearby non-agricultural areas.

Sampling and data quality considerations

The DHS data have been widely used in health and mortality studies (IRD-Macro Intl, 1991); nevertheless, DHS are broad-scale and cross-sectional, not customized, a priori, to a given research question. As a result, the sampling frame or set of questions on a particular topic may have some shortcomings for the question at hand. In this study, two such issues arose.

All children were selected for the survey by virtue of being born to women who were resident in surveyed households. It is important to note that children not living in the same household as their biological mothers, including orphans and foster children, were not surveyed. This is a concern since orphans and foster children, because they live either in institutional settings or other family arrangements that may already be vulnerable, may receive fewer nutritional, educational and health investments than children who reside in households with their biological parents. In a recent study of Rwanda, Siaens et al. (2003) found significant differences between orphans and non-orphans in terms of school enrolment, child labour and malnutrition rates, among other outcomes.

Some recent DHS, including those carried out since 1999 in Ethiopia (CSA, ORC Macro, 2001), Mali (CPS-MS, DNSI, ORC Macro, 2002) and Namibia (MOHSS, 2003) and two others not included in this study population,[10] measured all children resident in surveyed households, but most surveys in this sample did not. Namibia is the only country among these where children not living with their mothers have significantly lower weight-for-age Z scores than children living with their mothers. Because Namibia has the greatest percentage of such children among this set of countries, it is also possible that this is a sample-size effect.[11] Thus, while the potential for bias is noted, significant findings

[10] Eritrea (NSEO-ORC Macro, 2003) and Zambia (CSO, CBH, ORC Macro, 2003) are not included in this portion of the study.

[11] AIDS has been documented to cause a wide range of development challenges (Sachs, 2005), in part through its increase of fragmented homes. HIV/AIDS prevalence is significantly higher in Africa than anywhere else in the world, perhaps acting as incentive for the DHS to include children who do not reside with their biological mother (although they do not include children residing in institutions). Of these five countries, Zimbabwe (33.7%), Namibia (22.5%) and Zambia (15.6%) have high prevalence of HIV, but Eritrea (2.8%) and Mali (1.7%) do not (World Bank, 2004).

Table 3
Category variables used in Phase 1, individual analysis of underweight children

Variables	Rural only (% of sample)	Urban and rural (% of sample)
Country of origin		
Burkina Faso (reference category)	2.63	2.39
Benin	1.18	1.39
Central African Republic	0.95	1.29
Côte d'Ivoire	4.26	5.29
Cameroon	4.43	4.83
Ethiopia	18.24	16.09
Ghana	3.78	4.11
Guinea	1.08	1.25
Kenya	10.96	10.43
Madagascar	6.05	6.05
Mali	2.62	2.76
Malawi	3.26	2.97
Nigeria	21.59	22.18
Niger	5.07	4.79
Namibia	0.34	0.39
Chad	1.81	1.80
Togo	1.59	1.63
Tanzania	8.09	7.95
Zimbabwe	2.09	2.40
Family characteristics		
First born	18.23	19.45
High birth order	38.91	36.49
Twin	1.98	2.06
Male	50.05	50.21
Mother is head of household	8.83	8.78
Maternal schooling		
No schooling	57.15	52.19
Attended primary school	33.05	33.41
Attended secondary school	9.80	14.40
Mother is unemployed	28.33	29.32
Mother has skilled employment	4.08	4.80
Household characteristics		
Electricity	5.54	16.37
Radio	43.20	49.41
Television	4.11	10.69
Well water	49.36	43.09
Surface water	37.86	31.57
Flush toilet	0.76	4.67
No toilet	52.70	44.44
Household has finished floor	24.12	34.03
Child has been fully vaccinated by age 1	37.34	42.01
Child had a fever in the past 2 weeks	40.37	39.21
Length of growing period		
Between 120 and 299 days	65.51	65.99
Too short (fewer than 120 days)	14.54	13.60
Too long (more than 299 days)	19.94	20.42
Urban residence		20.67

Table 4
Continuous variables used in Phase 1, individual analysis of underweight children

Variables	Rural only			Urban and rural		
	Min	Max	Mean	Min	Max	Mean
Family characteristics						
Birth weight	1	5	3.15	1	5	3.18
Percentage of household under age 5	0	75	32.23	0	75	32.02
Child's age (months)	12	35	22.85	12	35	22.84
Mother's age	15	49	28.49	15	49	28.36
Children ever born to mother	1	18	4.25	1	18	4.11
Percentage of mother's children that died	0	86	12.38	0	86	11.37
Land characteristics						
Sandy soil	0	100	10.23	0	100	10.15
Aluminium toxicity	0	100	21.02	0	100	21.40
Low nutrient reserves	0	100	34.32	0	100	34.89
Shallow soil	0	100	12.96	0	100	12.41
Soil type is organic	0	17	0.56	0	17	0.57
Elevation	0	3814	906.63	0	3814	850.18
Slope	0	1765	155.45	0	1765	141.84
Precipitation (average daily rainfall mm)	22	3165	1225.29	22	3165	1234.35
Per capita production within 25 km	0	44,311	77.98	0	44,311	80.52
Tree cover (% within 10 km)	0	70	16.60	0	70	16.01
Pasture (% within 10 km)	1	88	38	1	88	37.00
Cropland (% within 10 km)	0	84	16.85	0	84	16.20
Other						
Malaria (stable transmission index)	0	38	13.22	0	38	13.76
Population density, (measure)	0	3762	173.74	0	11440	427.55
Distance (km, straight line) to nearest port	2	1813	504.28	1	1813	477.92
Distance to nearest town of >25k persons	0	89	8.53	0	118	7.36

occur in only one of five countries. We recommend that future data be collected so as to test for this bias more completely.

Anthropometry data availability varies widely between countries. In the anthropometry sample in Guinea, 41% of children have no valid measurement data, while in Nepal this rate is only 4%. Children may lack valid data for several reasons: they could be absent from home at the time of the survey, their mothers could refuse permission for their measurement or some combination of height, weight and age values could be sufficiently implausible as to cast doubt on all measurements for the child. We assume no systematic bias is associated with these omissions.

A child's underweight status reflects both short- and long-term effects of inadequate dietary intake and the effects of infectious diseases, short-term shocks such as floods, droughts and conflict or a good rainfall year, as well as longer-term political and economic stability at the time that DHS were conducted. The environmental variables used in this analysis represent long-term or predominating conditions, rather than short-term fluctuations. Some of these variables are known to have considerable temporal variation (e.g., climate) whereas others (e.g., soils, elevation and slope) are not expected to change markedly over the short or medium term. Although the DHS rarely field a survey during times of crisis, they often implement one when a country is in a rebuilding period. Future studies should consider the impact of time-varying variables with known spatial specificity, namely climate variability and political conflict.

Methodology

We undertook a simple ordinary least squares (OLS) regression analysis to test hypotheses from the UNICEF framework. Incompatibility across surveys constrained the variables selected for analysis. Tables 3 and 4 describe the variables included in the analysis. The outcome variable used was Z scores of weight-for-age of children; we assigned children whose weight-for-age was more than 2 standard deviations (SD) above the mean a score of 2 SD above the mean in order to remove any impact of overweight children, for whom the model is not designed. We further restricted the analysis to children aged 1–3 because below that age, and particularly below 6 months, they may have more error in their anthropometry. Also, they had censored observations on some of the explanatory variables (e.g., vaccination rates, because children under age 1 are not required to be fully vaccinated based on WHO guidelines). Children above age 3 were not surveyed in some countries.

We inverted and rescaled the weight-for-age scores so that the least underweight child was assigned a score of 0 and the most underweight had a score just below 8. Thus, positive covariates in the regression correspond to high shares of underweight children and negative covariates are interpreted as being associated with lower shares of underweight children. We tested results for this scaled index against results for the raw Z scores, and they showed no significant difference, thus, the regression results reported here use the scaled index because they are easier to interpret. In addition, we tested the models using height-for-age Z scores (scaled and unscaled). These results are summarized in the text but not in the tables.

All households in a given EA have the same values of georeferenced variables. Because of this and the complex sampling frame, we used OLS regression with a cluster flag in Stata to account for possible spatial error.[12] We ran various models for hypothesis testing whereby poorly performing, and then ultimately potentially endogenous variables are removed from the final equation.

Results

Table 5 shows the results of these regressions. Models 0–6 were conducted using only rural children to maximize the potential impact of some of the agro-climatic variables. Model 0 shows the differences among underweight children by country. With Burkina Faso as the reference category, Chad, Mali and Madagascar experience the same levels as do children in Burkina Faso, and children in Zimbabwe experience the most favourable relative circumstances. Countries are again controlled for in Model 6 and much of the variation in the underspecified Model 0 will be accounted for by variation in socio-economic, demographic and spatial factors that differ across the countries of the study region.

[12] The technique was successfully used elsewhere (Balk et al., 2004) and accounts for the complex sampling frame. This model specification is a first-order attempt to account for cluster-specific spatial autocorrelation but a more formal spatial regression is not specified here, as it is the subject of ongoing work. At this scale, the specification of spatial regression has proven challenging to implement in the existing spatial regression software packages.

The variables introduced in Model 1 are basic demographic characteristics of each child and, except for child age, are exogenous to the child but are necessarily intended as more than control variables. Older children are less likely to be underweight (Madise et al., 1999). However, unlike previous work, quadratic or logarithmic relationships (tested in models not shown) were not significant. We expected children born as a multiple birth (e.g., twins or triplets) to be more underweight, even in this age range, given the strong association exhibited by this variable on childhood mortality in the region (Balk et al., 2004). While we found no significance for this variable in Model 1, we retained it in later models, in which it becomes evident that twins tend to be more underweight than other children. Similarly, children in high parity births, who may also need to compete disproportionately for resources, are more likely to be underweight. We expected that male children would be less underweight than female children because evidence shows that daughters receive fewer investments on average and therefore we expected they would receive fewer nutritional ones as well. However, the regression results suggest just the opposite, reiterating surprisingly widespread evidence from other studies (Madise et al., 1999; Charmarbagwala et al., 2004). Nutrition experts also have suggested that such differences may be hard to observe (Kativa Sethuraman, personal communication, December 2003).[13] This effect persists throughout the models. Lastly, the proportion of the household that is under age 5 in a household positively correlates with their underweight status. Two potential explanations are likely. More children can translate into more competition for food, and fewer adults per child means less labour supply.

In our initial Model 1, we included a term for the duration of breast-feeding, expecting that breast-feeding would exhibit a strong protective effect against children being underweight, as has been demonstrated widely in the literature (Van Landingham et al., 1991; Menken and Kuhn, 1996). However, we observed the contrary: the longer a mother breast-feeds her child, we found, the greater the likelihood that the child is underweight. We determined that this effect was likely due to endogeneity, a finding with which other studies have also had to come to terms. Engle et al. (1999) find, for example, that education increases the mother's opportunity cost and reduces her duration of breast-feeding. Further, to the extent that long-duration breast-feeders are poorer than other women (Osinusi, 1987; Perez-Escamilla et al., 1999), they may not be able to provide the same supplemental foods to their children as other women, and themselves may be poorly nourished, leading to lower-quality breast milk. Thus, breast-feeding was not accounted for in models reported in Table 5.

Model 2 includes maternal characteristics. Increases in mother's age are inversely related to underweight status. One interpretation of this is that experience matters, although whether that is specifically maternal experience is not clear. The effect is marginally more important at younger ages, as expected, with a logarithmic specification performing marginally better than a linear one. The children born to mothers who have experienced high child mortality of their own children have much greater chances of being underweight. Education, both primary and secondary, exhibits strong downward pressure on being underweight.

[13] Garg and Morduch (1998) suggest that gender effects are observed with birth interval, and gender of succeeding births is accounted for, whereby sisters of brothers fare comparably worse than sisters of sisters.

Table 5
Child-level regression coefficients predicting underweight status

	Models								
	0	1	2	3	4	5	6	7 (6 + urban)	8 (stable)
Country dummies (ref. = Burkina Faso)									
Benin	−0.494**						−0.211**	−0.214**	−0.229**
Central African Republic	−0.363**						−0.218*	−0.154†	−0.196**
Côte d'Ivoire	−0.464**						−0.155†	−0.191**	−0.209**
Cameroon	−0.711**						−0.273**	−0.245**	−0.274**
Ethiopia	0.163**						0.065	0.035	0.157**
Ghana	−0.452**						−0.044	−0.027	0.011
Guinea	−0.541**						−0.211*	−0.158*	−0.160*
Kenya	−0.681**						−0.500**	−0.491**	−0.402**
Madagascar	0.030						0.309**	0.332**	0.314**
Mali	−0.026						−0.067	−0.107†	−0.132*
Malawi	−0.467**						−0.281**	−0.288**	−0.284**
Nigeria	−0.145†						0.009	0.075	0.070
Niger	0.412**						−0.086	−0.065	−0.154†
Namibia	−0.540**						−0.434**	−0.359**	−0.438**
Chad	0.042						−0.197**	−0.165**	−0.217**
Togo	−0.410**						−0.183*	−0.184**	−0.188**
Tanzania	−0.263**						−0.128	−0.097	−0.032
Zimbabwe	−0.959**						−0.638**	−0.569**	−0.560**
First born		−0.051							
High birth order		0.138**					0.121*	0.124**	0.126**
Twin		0.122					0.268†	0.324**	0.333**
Birth weight		−0.215**					−0.157**	−0.172**	−0.172**
Sex		−0.124**					−0.109**	−0.114**	−0.112**
Percentage of household under age 5		0.005**					0.179	0.174†	0.176†
Mother is head of household		−0.079†					0.105**	0.070*	0.077*
Child's age (months)		−0.007**					−0.004*	−0.004*	−0.004*
Mother's age (log)			−0.370**				−0.203*	−0.254**	−0.255**

Maternal schooling (ref. = none)						
Attended primary school	−0.377**			−0.104**	−0.094**	−0.103**
Attended secondary school	−0.698**			−0.308**	−0.318**	−0.330**
Mother is unemployed	0.111**			0.065*	0.047†	0.050†
Mother has skilled employment	0.057					
Children ever born to mother	0.015†					
Percentage of mother's children that died	0.004**			0.202*	0.173*	0.181*
Electricity		−0.203*		−0.142*	−0.130**	−0.132**
Radio		−0.228**		−0.123**	−0.120**	−0.122**
Television		−0.107				
Water/sanitation principal component 1		0.042**		−0.017	−0.008	
Water/sanitation principal component 2		0.042*		−0.014	0.006	
Household has finished floor		−0.205**		−0.091*	−0.102**	−0.103**
Child has been fully vaccinated by age 1		−0.285**		−0.064*	−0.067**	−0.072**
Child had a fever in past 2 weeks		0.201**		0.202**	0.181**	0.182**
Soil characteristics						
Sandy			0.004**	0.003*	0.002*	0.002*
Aluminium toxicity			0.001			
Low nutrient reserves			−0.002†	0.000	0.000	
Shallow			0.000			
Type is organic			−0.068**	−0.016†	−0.016*	
Length of growing period (ref. = 120–299 days)						
Too short (<120 days)			0.275**	0.087	0.103†	0.128*
Too long (>299 days)			0.101*	0.115*	0.093*	0.096*
Elevation (1000 m)			0.099**	0.002	0.020	
Slope			0.000			
Precipitation (average daily rainfall, mm)			0.036			
Per capita production within 25 km			0.000			

(*continued on next page*)

Table 5 (*continued*)

	Models								
	0	1	2	3	4	5	6	7 (6 + urban)	8 (stable)
Tree cover (% within 10 km)					−0.008**		−0.003†	−0.004**	−0.004**
Pasture (% within 10 km)					0.003*		0.322	0.251	0.365*
Cropland (% within 10 km)					0.005**		0.005**	0.005**	
Malaria (stable transmission index)						0.002			
Population density (1000 people/km^2)						−0.151†	−0.071	−0.003	
Distance (1000 km straight line) to nearest port						0.571**	0.303**	0.253**	0.265**
Distance (km) to nearest town of >25 K persons						0.001			
Urban residence								−0.136**	−0.143**
Constant	4.045**	4.685**	5.111**	4.034**	3.620**	3.555**	4.969**	5.217**	5.219**
R^2	0.075	0.040	0.053	0.061	0.068	0.022	0.154	0.178	0.176
F	68.91	31.29	46.54	55.63	33.82	34.2	42.65	60.70	69.51
Number of observations	25,674	25,448	25,674	24,882	25,667	25,671	24,673	33,655	33,658

Significant at †$p < 0.1$, *$p < 0.05$ and **$p < 0.01$.

Model 3 controls for household-level socio-economic characteristics, and those representing health investments. Children in households that were electrified and had radios were less likely to be underweight, not surprisingly, although we observed no effect based on television ownership. We found similar downward pressure on underweight status in homes with finished floors. Because of high correlation between data on water and sanitation, we combined sources into two indices using principal components. The first component is correlated positively with well water and negatively with surface water. The second is correlated positively with the complete lack of toilet facilities and negatively with flush toilets. Both were significant in the direction expected. However, neither is significant in fuller models.

Children who were fully vaccinated by age 1 were less likely to be underweight, presumably both because they are healthier and because they have parents who are more proactive in terms of direct health investments. Analogously, children who had experienced recent fever were more likely to be underweight. A recent fever is assumed as indication of the child's more general health status, although there are no other ancillary data in the data set to confirm this.

Model 4 introduces soil fertility constraints and other environmental variables. We tested four variables describing physical and chemical soil characteristics in this model as well as whether the soil type was organic (i.e., characterized by wetness, low bulk density and low fertility, particularly nitrogen and micro-nutrient deficiencies). These soil variables were considered to be the most distinctively difficult for food production (Pedro Sanchez, personal communication, May 2004). Children living in areas with soils that have low nutrient reserves or are sandy are more likely to be underweight, with a much stronger effect being observed on sandy soils. Additionally, as we expected, soils with high organic content were found to be associated with lower levels of underweight children.

We also included length of growing period (LGP) in the model. Children living in areas with growing periods that were either very short (fewer than 120 days) or very long (more than 299 days) were more likely to be underweight than children living in areas with more optimal growing periods, consistent with our expectations. Agro-ecological zones with very short or very long LGP (due to rot) do not produce enough food; therefore, agricultural communities living in these areas may face food shortages and malnutrition.

Elevation is also associated with being underweight but average terrain slope showed no effect. Children living in areas with more tree coverage were less likely to be underweight but, counter to our a priori expectations, children living near pasture and cropland were more likely to be underweight. We know that cultivated areas tend to be more densely populated than other rural areas (McGrahanan et al., in press), a variable that is independently included in subsequent models. We asked whether farming systems categorically differ in their socio-economic characteristics, for example, whether educated families are less likely to live in the most vulnerable systems (either because they have migrated to less vulnerable ones, or because those regions have fewer schooling opportunities). Although there are statistically significant differences in mother's educational level by farming system (not shown), mothers with the highest average education were found in a wide range of farming systems (e.g., both rice-tree cropping and sparse agriculture) dismissing concern over potential self-selection.

Aside from these direct agro-climatic conditions, a few additional spatial variables are entered in Model 5. These include an index of the stability of malaria transmission risks (Kiszewski et al., 2004), population density and distance to the nearest port and city of

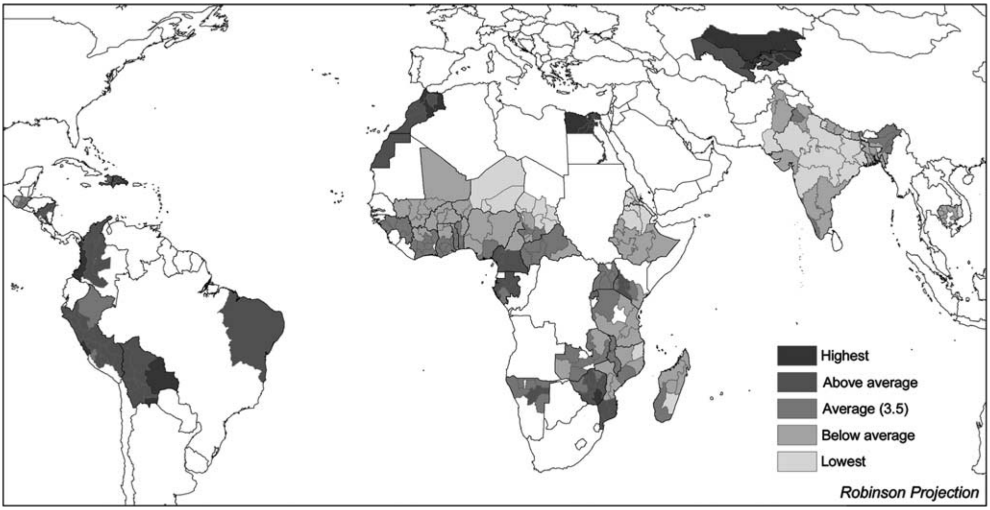

Fig. 2. Average regional weight-for-age Z scores, actual values.

25,000 persons or more. Of these, only population density (weakly) and nearness to port (strongly) were associated with not being underweight. We assume that the population density captures a large degree of urbanness and thus reduces the likelihood of an independent effect of nearness to an urban centre. The malaria transmission index showed no effect. The coarse (half-degree) resolution of the data may have reduced any potential effect.[14]

Model 6 enters all variables that showed an effect significant at the 0.1 level in earlier models (including some of those discussed as flawed or with unexpected effects above) as well as all the country dummies. While some variables are no longer significant (e.g., elevation), most are simply dampened by the complexity of the model. Once controlling for many covariates, nearly all the country dummies exhibit weaker and less significant effects.

The Model 7 regression was run on a fuller subsample of the data – urban and rural children in the same age range. In this model, a dummy variable for urban dweller was included. As expected, urban children are less likely to be underweight than their rural counterparts. All other significant regressors from Model 6, even the soil modifiers, remain significant (some more, others less) in Model 7.

In preparation for Phase II of this study, we generated Model 8, in which endogenous and unexplainable variation (as well as variables that would be conceptually meaningless at the meso-scale) was removed from Model 7. This has the overall impact of slightly reducing the significance of the overall model (i.e., the R^2, or explainable variance term) but provides a stable and unbiased estimate of underweight children at the micro-level for comparison with the meso-scale analysis.

[14] An analysis of mortality had similar findings showing a strong bivariate relationship with the malaria index but a significant relationship in the multivariate model was not sustained, partly due to the strong co-linearity with other country-level control variables (Balk et al., 2004) due to the coarseness of the index. Future studies should use more highly resolved data, if and when they become available.

Using height-for-age Z score as the dependent variable instead of weight-for-age produced similar but weaker results in the eight models (not shown). The only factors for which the direction of a significant effect changes are country dummies. In general, fewer effects are significant and more weakly so. Mother's primary education, proximity to ports, vaccinations, tree cover and mothers heading households all lose much or all of their significance in most or all of the models in which they appear. Multiple births, however, are more significantly correlated with low height-for-age scores. The effects of elevation and the malaria transmission index also become significant, although in the unexpected direction.

To summarize the child-level analysis, numerous factors emerged as clearly significant. In general, those operating at the household level showed clearer effects than those at the environmental level. The specification of this model and its associated data, however, did not allow us to determine whether households are effective mediators of environmental risk. For example, elevation has no effect when household socio-economic variables are included, presumably because most of those with significant resources have been able to move to more favourable (i.e., lower) elevations. It is also plausible that there is far less measurement error in the household variables for which the survey instruments were designed than in the geospatial variables, which were originally measured for different units and purposes and associated with the surveyed regions through a complex and imperfect process. If the latter were true, the geospatial variables would perform better in a meso-level analysis where such imprecision can be tolerated.[15]

Phase II. Meso-level analysis of determinants of underweight children in Africa, Asia and Latin America

In this phase, we aggregated variables for all children to the subnational survey region in which they lived and then undertook regressions. Because we did not require precise spatial information on the location of the household, this analysis included data from countries that did not have the EA location information needed for the analysis in Phase I. Thus, the Phase II analysis incorporates regions in several countries in Asia and Latin America as well as additional countries in Africa. To make the results as comparable as possible to those from Phase I, we only worked with rural households.[16] To reduce standard errors, we omitted subnational regions where fewer than 50 children have valid data. This left 319 subnational units, drawn from 45 countries (Fig. 2). Because we are not attributing regional-level associations to individual-level phenomena but rather characterizing relationships strictly at the regional level, the ecological fallacy is not a critical problem (Schwartz, 1994; Diez Roux, 2004).

Data

In this phase, the outcome variable is the average weight-for-age Z score among children aged 1–3 in each region. As in Phase I, children who were more than 2 SD above

[15] Because so few differences emerge between the wasting and stunting analyses, we did not conduct a meso-scale analysis using stunting as the malnutrition indicator.

[16] This is not to suggest that urban malnutrition is not an important issue, rather just one that cannot be as well specified in the meso-scale analysis as it could have been in the individual-level analysis with more geographic information.

Table 6
Aggregate-level regression results predicting underweight conditions in subnational regions

	Model									
	0	1	2	3	4	5	6	7	8	9
Country dummies (ref. = India)										
Bangladesh	0.222†							−0.090		0.065
Benin	−0.415**							−0.446*		−0.166
Bolivia	−1.016**							−0.935**		−0.527**
Brazil	−1.158**							−0.865**		−0.610*
Burkina Faso	0.095							−0.208		0.157
Cambodia	0.092							−0.034		0.169
Cameroon	−0.807**							−0.577**		−0.338*
Central African Republic	−0.278*							−0.629**		−0.263†
Chad	0.195*							−0.419		−0.129
Colombia	−1.116**							−0.286		−0.140
Comoros	−0.456*							−0.170		−0.376
Côte d’Ivoire	−0.348**							−0.763**		0.055
Dominican Republic	−1.465**							−1.168**		−0.732**
Egypt	−1.751**							−0.072		−1.029**
Eritrea	0.345*							−0.455**		0.011
Ethiopia	0.177†							−0.360†		−0.268*
Gabon	−0.825**							−0.376†		−0.099
Ghana	−0.321**							0.151		−0.035
Guatemala	−0.267*							−0.630**		0.118
Guinea	−0.460**							−0.645**		−0.451*
Haiti	−0.823**							−0.718**		−0.718**
Kazakhstan	−1.583**							−1.824**		−1.305**
Kenya	−0.601**							−0.473**		−0.197
Kyrgyz Republic	−1.199**							−1.374**		−0.840**
Madagascar	0.020							−0.073		0.132
Malawi	−0.474**							−0.553**		−0.237
Mali	0.050							−0.367†		−0.003
Morocco	−1.156**							−0.912**		−0.751**
Mozambique	−0.331**							−0.519**		−0.287*
Namibia	−0.503**							−0.617**		−0.115
Nepal	0.211							0.034		0.066

Nicaragua	−1.045**					−0.734**		−0.606**
Niger	0.448**					−0.286		0.030
Nigeria	−0.055					−0.136		0.161
Pakistan	0.031					−0.299†		−0.223
Peru	−0.965**					−0.591**		−0.434**
Rwanda	−0.454**					−0.514**		−0.388**
Tanzania	−0.162					−0.167		0.095
Togo	−0.330*					−0.342		−0.018
Uganda	−0.500**					−0.656**		−0.340*
Uzbekistan	−1.095**					−1.111**		−0.645**
Zambia	−0.352**					−0.550**		−0.146
Zimbabwe	−0.860**					−0.783**		−0.422**
HH size factor 1		−0.047				−0.091**	−0.078**	−0.084**
HH size factor 2		−0.206 **				−0.055*	−0.074**	−0.041†
Fraction mothers head of HH		−0.660				−0.174		
High-fertility index		0.295**				0.057		
Average mothers age			−0.078**			−0.010	−0.052**	−0.007
Fraction mothers with primary education			−0.841**			−0.598**	−0.734**	−0.698**
Fraction mothers with secondary education			−1.311**			−0.204	−0.520**	−0.567**
Fraction mothers in skilled employment			−0.007			0.357		
Fraction mothers unemployed			−0.665**			−0.091		
Asset index 1				−0.288**		−0.138*	−0.324**	−0.131*
Asset index 2				−0.359**		−0.190**	−0.342**	−0.237**
Fraction children fully vaccinated				−0.408**		−0.070		
Fraction children with fever near survey				0.220		0.063		
Sandy soil (%)					0.012**	0.003	0.004*	0.003†
Territory with growing season >300 days (%)					−0.271*	−0.099		
Territory with growing season <120 days (%)					−0.608**	0.013		

(*continued on next page*)

Table 6 (*continued*)

	Model									
	0	1	2	3	4	5	6	7	8	9
Fraction of territory with trees					−0.010**			0.000		
Average elevation (1000 m)					−0.003			0.089†		
Average slope					0.007			−0.028*		
Average food production (kcal per capita)					−0.005**			0.002		
Average Malaria Index						0.019**		0.002	0.011**	0.000
Pop density						0.546†		0.087	0.176*	0.121
Average distance to nearest port						0.020*		0.016		
Fraction of territory occupied by small urban areas						0.015**			0.005*	0.001
Fraction of territory occupied by medium urban areas								−0.001		
Fraction of territory occupied by large urban areas								0.000		
Fraction of territory within 2 km of road						−0.843*		−0.219		
Fraction of territory within 15 km of road						0.356		0.023		
Fraction of territory near railroad						−1.729†		−0.444		
Density factor 1 (high density, lots of small urban areas)						−0.126		0.015		
Density factor 2 (large urban areas)						−0.117				
Political conflicts since 1975 (low intensity)							0.043†	0.013		
Political conflicts since 1975 (medium intensity)							−0.003	0.015	0.019**	−0.006
Political conflicts since 1975 (high intensity)							0.020*	0.009		
Political conflicts during 1990s (low intensity)							−0.068*	−0.033		
Political conflicts during 1990s (medium intensity)							0.001	−0.050		
Political conflicts during 1990s (high intensity)							−0.068†	−0.026		
(Constant)	3.980**	3.518**	6.521**	3.570**	3.922**	3.084**	3.461**	4.491**	5.069**	4.249**
R^2	0.831	0.119	0.369	0.638	0.265	0.233	0.033	0.903	0.769	0.889
F	31.38	10.63	36.67	138.5	15.63	10.15	1.73	27.77	82.08	37.08

Significant at $^{\dagger}p < 0.1$, $^{*}p < 0.05$ and $^{**}p < 0.01$.

the mean were re-coded; and the scores were inverted, so that high scores corresponded to highly underweight conditions, and adjusted arithmetically so that their lowest value was zero. Thus, for example, a score of 0 corresponds to 2 SD or more above the mean, and a score of 3 would correspond to 1 SD below the mean. The outcome variable for the subnational units in the study ranged from 1.4 to 5.0, with a mean of 3.5. This outcome variable correlates quite highly with the percentage of all children who are more than 2 SD below the mean (i.e., Pearson's correlation coefficient is 0.96).

We calculated subnational aggregations of all the relevant variables tested in Phase I. For example, the Phase I variables that indicate the mother's educational status were aggregated to two variables that indicate the percentage of mothers in the region with some primary education and the percentage with some secondary education. For these educational variables, the aggregated versions are directly analogous to the household-level counterparts. In other cases, however, the variables need to be interpreted in a different light when aggregated. Birth order, for example, serves as a child-level control variable in the Phase I study. A child who enters the family as the fifth child or higher is significantly more likely to be underweight than a child who is not. At the subnational level, the aggregation of this variable serves not as a control variable for individual risk but rather an indicator of fertility levels within the region.

We found that, for the subnational aggregates, measures that were substantively related often performed poorly when entered into a regression together. This was particularly so for the household asset measures; although each was separately negatively correlated with underweight outcomes, when combined some became positively correlated. We therefore investigated principal components for functionally related variables and used factor scores in lieu of individual variables in some cases. We calculated two household asset indicators using principal component analysis. We further found that many of the control variables in the Phase I analysis served essentially as indicators of household size and composition at the regional level, and thus calculated two principal components, which we designated HHSize1 and HHSize2. Lastly, we calculated a high-fertility index by using the factor loads from two variables, the percentage of children who are first-born and the percentage of children who are fifth-born or higher. We opted for this rather than for measuring the total children ever born to the mother to account a bit more for distributional effect than simply the average total fertility in the region.

Methodology

We sought to explain the variation in the underweight measure using a sequence of OLS regression models emphasizing different sets of drivers in the same manner as Phase I, except that here we treated each subnational region as an independent observation. We then combined the drivers into a single model and finally estimated a model with only statistically significant factors. This latter model constitutes a composite risk model, whereas the others identify risks associated with specific factors.

Results

Table 6 summarizes results. Model 0 shows the degree to which national dummy variables can account for differences at the level of the subnational unit. Largely because the average country has only 7.25 subnational units, the country dummies perform quite well in this model, accounting for 83% of the variance.

Model 1 considers the effect of regional demographic structural factors. Regions with large households have fewer underweight children. This is consistent with findings from the economic demography literature showing that large households are better equipped to pool risk than are smaller households (e.g., Townsend, 1994). We also find here that regions with high fertility have more underweight children; again this is consistent with prevailing understanding. Altogether these factors account for 12% of the variance.

In Model 2, maternal attributes are explored. Education plays a major impact, as expected. The proportion of mothers with some primary education reduces underweight status, as does the proportion with some secondary education. We find that the average age of mothers is negatively associated with underweight status, which is consistent with Phase I findings. In this model, however, maternal unemployment unexpectedly also reduces underweight status. Whether this indicates that an underlying factor associated with unemployed mothers – such as regional economic factors, i.e., women may not work in wealthier regions – has an effect, requires more investigation. Model 2 accounts for 37% of the variance.

In Model 3, we explore the role of household-level investments. We find a strong effect for the physical assets and a significant effect from vaccination rate. However, the fever measure, which was significant as a child-level control in Phase I, does not perform well as a regional aggregate. Consistent with the conclusion that tangible household assets are a strong determinant of life prospects, this model accounts for 64% of the variance. Because of the strong association, we map the risk factor derived from these measures (Fig. 3).

Model 4 examines the impact of physical factors relevant to agricultural production. Because of complications in processing the soil data, we only tested one soil constraint in Phase II but would have liked to test more. We found that the percentage of a region with sandy soil is positively associated with underweight status. The tree-cover measure also performed here as in Phase I. The growing season measures that performed as hypothesized in Phase I had the opposite effect in Phase II; regions with very long or very short growing seasons had lower underweight problems than other regions. In future tests, we would want to calculate population-weighted aggregates of growing season as our results here may be a function of territorial averages that do not reflect the conditions

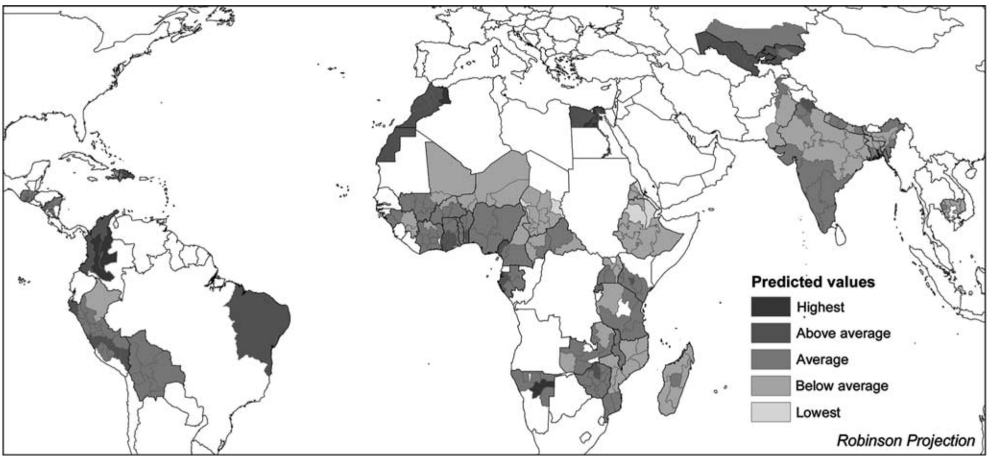

Fig. 3. Predicted values of regional weight-for-age Z scores, wealth and health model.

under which most people actually live. For example, a region with a wet and a dry zone will most likely have most of its people living in the wet zone; therefore, it would be misleading to include the dry zone's lower growing season as a contributor to the well-being of the people living in the wet zone. This model altogether accounts for 26% of the variance.

Model 5 looks at other physical characteristics of the region. We find significant relationships in the expected direction for exposure to malaria, distance from ports and density of road networks and railroads. For overall population density, we find that high-density levels are associated with higher levels of underweight status; normally urban children are less underweight than rural children, which is borne out by the Phase I results. Additional sensitivity analysis is called for here. We do not yet know how consistent the pattern is across regions or whether a small number of cases, such as the high underweight values in India that are accompanied by high population densities, are masking a different pattern elsewhere. Regions with a large number of small urban areas have higher underweight status; perhaps because such regions are more likely to be experiencing settlement growth that is outpacing their ability to invest in necessary infrastructure, thereby placing their populations at greater risk. Model 5 accounts for 23% of the variance.

In Model 6, we explore the effect of political conflict in a region, using a gridded database of conflict that we created using the Uppsala Conflict Database (Gleditsch et al., 2002). Based on our work with a data set of global subnational infant mortality rates, we expected these variables to be significant (Levy et al., 2003); however, they did not perform well in this test. In the results reported here, some of the variables are significant but half the coefficients have the opposite sign expected (which is probably a function in part of the high collinearity of these measures). We find that the single best of these measures in accounting for underweight status is the cumulative number of medium-intensity conflicts. However, the effect of conflicts since 1975 is weak. This model overall accounts for 3% of the variance.

Model 7 includes all the variables used in each of the above models. It accounts for 90% of the variance in underweight status. As in the Phase I results, most of the country dummies remain significant. The additional variables that remain significant in this model are the two household size factors, mother's primary education, the two asset indices, and slope and elevation. The importance of these variables, because they are highly significant even when controlling for everything else and when including country dummies, is a robust finding.

Our suspicion is that including country dummies obscures some of the effect of the geospatial variables because some of the effects of these are exerted on the country as a whole as well as on the specific subnational regions. Countries with high malaria exposure, for example, have higher underweight status. There is an endogenous dynamic in which spatial drivers such as this exert a systemic effect on a country that in turn redounds to the detriment of regions and households, making it hard to discern precise signals in a static analysis of this sort. We cannot know in which cases the impact of the country dummies in bumping some variables out of significance reflects this kind of underlying dynamic and when it does not. Under the circumstances, we estimated Model 8, which removes the country dummies and retains only the remaining statistically significant variables. In this exercise, sandy soil, malaria exposure, population density, small urban areas and medium-intensity political conflicts all become significant. The non-spatial variables that are significant in this model include the two household size indicators, mother's age, both education measures and the two asset indicators. In some ways this model represents our best-informed guess about the relevant risk factors faced by rural children in developing countries. Even without the country dummies, the amount of variation accounted for by the model is high (77%). For the sake of completeness, we take the variables from Model 8, bring back all the country dummies and report the results as Model 9.

Conclusions

In this study, we attempted to capture the effects of geographic and environmental variables on child hunger, cross-nationally, at the household and regional levels of analysis. The strength of this approach was that such factors could be evaluated along with more commonly found predictors of child hunger but a disadvantage was that we were unable to include a full suite of individual- and household-level variables. Future studies, either by narrowing the geographic extent, or by improving the coverage of cross-nationally available data, should attempt to include a richer set of individual-, household- and regional-level factors in the same model.

Nevertheless, we were able to pinpoint a number of household characteristics that appear to be significant risk factors for child malnutrition. In particular, length of breast-feeding was inversely related to child nutritional outcomes, we argued due to a set of factors some of which we were unable to measure in this study, such as supplementary feeding. The UN Millennium Project HTF advocates for breast-feeding until age 2, and this study calls for better data to evaluate supplementary feeding practices in combination with nursing. The risk factors are largely consistent with prior studies, but the study also highlights the importance of proximity to markets, as represented by ports or urban residence. As no effect was found for smaller cities, consistent with the recommendation of the HTF it is likely that these towns and rural settlements may need extra investments of infrastructure, markets and so forth.

From a policy and research perspective it is noteworthy that the best-fitting models account for no more than 17% of the observed variance when all these factors are examined in the household-level models. Perhaps this is not surprising for a cross-national, multilevel analysis of survey outcomes that also includes externally observed environmental and geographic factors. It is consistent with the proposition that there are interactions between the factors we observe with some precision (household size, select household assets, and so on) and factors we observe with far less precision, perhaps with complex interactions and non-linearities, or observe not at all. As some have suggested, factors such as soil fertility, access to water, access to markets, availability of adequate and affordable medical care and other spatial variables may influence patterns of poverty and hunger, but the evidence for such influences will be hard to observe through a static comparison of individual families or children. Some small-scale studies have found strong effects from these kinds of relationships, and indeed we find that children living in regions with suboptimal agricultural growing seasons (i.e., too short or too long) or in selected farming systems, for example, were more likely to be underweight. Nevertheless, our effort to look for causal relationships using micro-data on a continental scale demonstrates that the empirical foundation for generalizing the relative impacts of such forces is significantly limited by an absence of adequate data, and potentially incompatibilities in the resolutions and units of analysis of the data that do exist.

At the regional level, the broad findings that emerge from the child-level study are supported on a more globally representative sample using subnational regions from more than twice as many countries. These models account for a greater percentage of the outcome's variance. This is consistent with our proposition that household and regional dynamics interact in complicated ways that are more difficult to discern at the individual level but which become somewhat clearer through a regional lens. Even in the analysis that omits the country-level dummy variables, the model's ability to account for variation in regional underweight status is high. In the micro-analysis, many important individual or contextual-level determinants are omitted, whereas individual variation is masked in the regional analysis (Diez Roux, 2004). One advantage of conducting an analysis at two scales is that policy recommendations could be qualified. That is to say, regions at high risk of malaria have higher shares of underweight children than in low risk areas. The fact that this variable was not significant at the household level does not imply that a household's individual risk of malaria is not important (and its associated consequences or co-determination of malnutrition), rather it suggests that improved specification of that risk is needed for a fuller understanding of behaviour and likely policy effectiveness.

The following factors emerge as significant in both methods and therefore we have high confidence that they play an important role in influencing the underweight status of children in developing countries: household composition and size, maternal education, household assets and low soil fertility (as measured by sandy soil constraint). The meso-scale analysis found that rural regions with small urban areas (as opposed to larger ones) were more likely to have higher malnutrition, thus, reinforcing the HTF emphasis on creating viable markets and access in predominantly agricultural areas. Risk of malaria also appears to be important, again reinforcing the importance of diseases that affect nutritional outcomes.

In both micro- and meso-scale analyses, more than any other block of variables, the influence of the spatial variables is much greater in the absence of the country-level dummy variables. One interpretation of this is that spatial variables provide substantive guidance to explain country-level variation. Since countries vary considerably in environmental and spatially explicit features, this method provides an entry to unlocking the ways in which these features affect health and poverty outcomes, rather than simply noting differences in country averages. The analysis shows the potential for constructing higher-risk zones that could be policy targets, although no such targeting was attempted herein.

The inability to measure the effect of agricultural productivity in this analysis was a notable concern, and our meagre results in this regard reflect primarily a need to invest in more appropriate data. Some progress is being made in this regard through the Agro-Maps project (http://www.fao.org/landandwater/agll/agromaps/interactive/index.jsp), but subnational data comparable across countries and dominant crops were not available at the time of this study. Many of the variables used were imperfect proxies; superior alternatives are in development and will shed better light on the impact such factors exert.

Finally, the differences within and between the results of micro- and meso-scale analysis clearly suggest that models which examine relationships where environmental and ecological factors are expected to be mediated by communities, households and individuals, need to refine both the data and models with which to tease apart interactions so that effective hunger and poverty reduction strategies can be implemented.

Acknowledgements

An earlier version of this article was prepared for the HTF of the United Nations Millennium Project. We thank Jordan Chamberlain, Yunnis Yi, Audra Query, Gabriela

Fighetti and Janina Franco for assistance in the preparation of data and some of the maps for this study. We thank Pedro Sanchez, Lisa Dreier and Nalan Yuksel for suggestions throughout the project. As part of the Millennium Project HTF's review process, we also received helpful comments from Peter Matlon, Christopher Barrett, Kostas Stamoulis and five anonymous reviewers.

References

Balk, D., Pullum, T., Storeygard, A., Greenwell, F., Neuman, M., 2004. A spatial analysis of childhood mortality in West Africa. Population, Space and Place 10 (3), 175–216.

Charmarbagwala, R., Ranger, N., Waddington, H., White, H., 2004. The Determinants of Child Health and Nutrition: A Meta-analysis. Department of Economics, University of Maryland and Operations Evaluation Department, World Bank, Washington, DC.

CIESIN, CIAT (Center for International Earth Science Information Network, Centro Internacional de Agricultura Tropical), 2003. Gridded Population of the World, version 3 (alpha). CIESIN, Columbia University, Palisades, New York. Available from: <http://sedac.ciesin.columbia.edu/gpw/>.

CIESIN, IFPRI, World Bank, CIAT (Center for International Earth Science Information Network, International Food Policy Research Institute, the World Bank, Centro Internacional de Agricultura Tropical), 2004. Global Rural–Urban Mapping Project (GRUMP): urban/rural population density grids (alpha). CIESIN, Columbia University, Palisades, New York. Available from: <http://sedac.ciesin.columbia.edu/gpw/>.

CPS-MS, DNSI, ORC Macro (Cellule de Planification et de Statistique du Ministère de la Santé, Direction Nationale de la Statistique et de l'Informatique, ORC Macro), 2002. Enquête démographique et de santé au Mali, 2001. Calverton, MD.

CSA, ORC Macro (Central Statistical Authority, Ethiopia, ORC Macro), 2001. Ethiopia demographic and health survey, 2000. Addis Ababa, Ethiopia and Calverton, MD.

CSO, CBH, ORC Macro (Central Statistical Office, Zambia, Central Board of Health, Zambia, ORC Macro), 2003. Zambia Demographic and Health Survey 2001–2002. Calverton, MD.

DeFries, R.S., Hansen, M.C., Townshend, J.R.G., Janetos, A.C., Loveland, T.R., 2000. A new global 1-km data set of percentage tree cover derived from remote sensing. Global Change Biology 6, 247–254.

Diez Roux, A.V., 2004. The study of group-level factors in epidemiology: rethinking variables, study designs, and analytical approaches. Epidemiologic Reviews 26, 104–111.

Dixon, J., Gulliver, A., Gibbon, D., 2001. Farming Systems and Poverty: Improving Farmers' Livelihoods in a Changing World. United Nations Food and Agriculture Organization, Rome. Available from: <http://www.fao.org/farmingsystems>.

Engle, P., Menon, P., Haddad, L., 1999. Care and nutrition: concepts and measurement. World Development 27 (8), 1309–1337.

FAO (Food and Agriculture Organization of the United Nations), 1995. Digital soil map of the world and derived soil properties, Rome, Italy.

FAO (Food and Agriculture Organization of the United Nations), 2003. Global map of irrigated areas, Version 2.1, Rome, Italy. Available from: <http://www.fao.org/ag/AGL/aglw/aquastat/irrigationmap/index.stm>.

Freedman, D.A., 2001. Ecological inference and the ecological fallacy. International Encyclopaedia of the Social and Behavioural Sciences 6, 4027–4030.

Foley, J.A., Costa, M.H., Delire, C., Ramankutty, N., Snyder, P., 2003. Green Surprise? How terrestrial ecosystems could affect earth's climate. Frontiers in Ecology and the Environment 1 (1), 38–44.

Garg, A., Morduch, J., 1998. Sibling rivalry and the gender gap: evidence from child health outcomes in Ghana. Population Economics 11 (4), 471–493.

Gleditsch, N.P., Wallensteen, P., Eriksson, M., Sollenberg, M., Strand, H., 2002. Armed conflict 1946–2001: a new dataset. Journal of Peace Research 39 (5), 615–637.

Graitcher, P.L., Gentry, E.M., 1981. Measuring children: one reference for all. Lancet ii, 297–299.

Habicht, J.P., Martorell, R., Yarbrough, C., Malina, R.M., Klein, R.E., 1974. Height and weight standards for preschool children. How relevant are ethnic differences in growth potential? Lancet 1 (7858), 611–614.

IIASA-FAO (International Institute for Applied Systems Analysis-Food and Agriculture Organization of the United Nations), 2000. Global Agro-Ecological Zones (GAEZ) 2000. CD-ROM, Rome, Italy.

IRD-Macro Intl (Institute for Resource Development-Macro International), 1991. In: Proceedings of the Demographic and Health Surveys World Conference, 5–7 August, 1991, Washington.
Kiszewski, A., Mellinger, A., Malaney, P., Spielman, A., Ehrlich, S., Sachs, J., 2004. A global index of the stability of malaria transmission. American Journal of Tropical Medicine and Hygiene 70 (5), 486–498.
Levy, M., Balk, D., Storeygard, A., Booma, G., 2003. Characterizing the global distribution of poverty. In: American Association of Geographers Annual Meeting, March, 2003, Philadelphia.
Macro International Inc., 1996. Sampling manual. DHS-III Basic Documentation No. 6, Calverton, MD.
Madise, N.J., Matthews, Z., Margetts, B., 1999. Heterogeneity of child nutritional status between households: a comparison of six sub-Saharan African countries. Population Studies 53 (3), 331–343.
McGrahanan et al., in press. Urban systems. In: Millennium Ecosystem Assessment: Condition and Trends, 5 vols. Island Press, Washington, DC.
Menken, J., Kuhn, R., 1996. Demographic effects of breastfeeding: fertility, mortality, and population growth. Food and Nutrition Bulletin 17 (4), 349–363.
Meybeck, M., Green, P., Vorosmarty, C., 2001. A new typology for mountains and other relief classes: an application to global continental water resources and population distribution. Mountain Research and Development 21 (1), 34–45.
MOHSS (Ministry of Health and Social Services, Namibia), 2003. Namibia Demographic and Health Survey 2000, Windhoek, Namibia.
NIMA (National Imagery and Mapping Agency), 1995. Performance specification: vector smart map (Vmap) level 0. MIL-PRF-89039, 9 February, Washington, DC.
NSEO-ORC Macro (National Statistics and Evaluation Office-ORC Macro), 2003. Eritrea Demographic and Health Survey 2002. National Statistics and Evaluation Office and ORC Macro, Calverton, MD.
Osinusi, K., 1987. A study of the pattern of breast feeding in Ibadan, Nigeria. Journal of Tropical Medicine and Hygiene 90 (6), 325–327.
Payne, P.R., 1990. Measuring malnutrition. IDS Bulletin 21 (3), 14–30.
Perez-Escamilla, R., Cobas, J., Balcazar, H., Holland Benin, M., 1999. Specifying the antecedents of breastfeeding duration in Peru through a structural equation model. Public Health Nutrition 2 (4), 461–467.
Rabus, B., Eineder, M., Roth, A., Bamler, R., 2003. The shuttle radar topography mission – a new class of digital elevation models acquired by spaceborne radar. ISPRS Journal of Photogrammetry and Remote Sensing 57 (4), 241–262.
Ramankutty, N., Foley, J.A., 1998. Characterizing patterns of global land use: an analysis of global croplands data. Global Biogeochemical Cycles 12 (4), 667–685.
Sachs, J., 2005. The End of Poverty. Penguin, New York.
Sanchez, P.A., 2002. Soil fertility and hunger in Africa. Science 295, 2019–2020.
Sanchez, P.A., Swaminathan, M.S., Dobie, P., Yuksel, N., et al., 2005. Halving hunger: it can be done. Final Report of the UN Millennium Project Task Force on Hunger, Earthscan, London.
Schwartz, S., 1994. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. American Journal of Public Health 84 (5), 819–824.
Siaens, C., Subbarao, K., Wodon, Q., 2003. Are Orphans Especially Vulnerable? Evidence from Rwanda. World Bank, Washington, DC.
Smith, L., Haddad, L., 2000. Explaining child malnutrition in developing countries: a cross-country analysis. Research Report 111, International Food Policy Research Institute, Washington, DC.
Tanner, J.M., 1981. A History of the Study of Human Growth. Cambridge University Press, Cambridge.
Townsend, R., 1994. Risk and insurance in village India. Econometrica 61 (3), 539–591.
UNICEF, 1990. Strategy for improved nutrition of children and women in developing countries, New York.
UNICEF, 1998. The state of the world's children, New York.
Van Landingham, M., Trussell, J., Grummer-Strawn, L., 1991. Contraceptive and health benefits of breastfeeding: a review of the recent evidence. International Family Planning Perspectives 17 (4), 131–136.
Whitehead, R.G., Paul, A.A., 1984. Growth charts and the assessment of infant feeding practices in the western world and in developing countries. Early Human Development 9, 187–207.
WHO (World Health Organization), 1995. Physical status: the use and interpretation of anthropometry. WHO Technical Report Series No. 854, Geneva.
World Bank, 2004. African development indicators, Washington, DC.