USAID FROM THE AMERICAN PEOPLE

ECONOMICS BRIEF

*Authors: Stephen A. O'Connell and Caroline Smith**

November 15, 2016

Economic growth is associated with reduced child undernutrition A comment on Vollmer *et al.* 2014

Abstract: Vollmer *et al.* (*Lancet Global Health*, 2014) employ Demographic and Health Survey (DHS) data to argue that there is virtually no empirical link between economic growth and early childhood undernutrition. We re-examine this link and come to a very different interpretation. The authors' own results imply a meaningful association between growth and undernutrition, and this link is much stronger once we correct for straightforward issues of measurement error, duration, and influence. An effective attack on early childhood undernutrition must be two-pronged, combining direct health interventions with vigorous efforts to advance economic growth.

Introduction

Does national economic growth improve the nutritional status of young children in low-income countries? Fundamental considerations would surely favor a positive impact, at least if growth is sustained. From the demand side, the health of children is both an intrinsic part of a household's well-being and a determinant of the household's current and future productivity. Even a rudimentary knowledge among parents of how food and other purchased inputs affect nutritional outcomes, and how these outcomes in turn affect the health of their children, would be expected to generate a link from increases in household income to improved diets or hygiene practices. An additional link would be expected to operate on the supply side, via government public-health expenditures financed by the tax revenues generated by growth.

There is no guarantee, of course, that these links are operative in particular cases. Armenia's national health system, for example, collapsed in the post-Soviet transition, with indicators of early childhood undernutrition continuing to deteriorate even as out-migration and favorable export prices finally drove a sharp recovery in real GDP per capita in the first half of the 2000s (Johnson 2007, Richardson 2013). Nor is there any guarantee that economy-wide growth will reach the households whose children are malnourished, especially if growth is transitory. How *large* the impact is from economy-wide growth to childhood undernutrition has therefore been a subject of empirical contention (Alderman *et al.* 2014).

In a recent contribution to this debate, Vollmer *et al.* (2014) use cross-country data from nationally-representative household surveys to investigate the association between economic growth and the nutritional status of children under the age of 3 in low- and middle-income countries. Their conclusion is striking:

> "In summary, the quantitatively very small to null association seen in our study suggests that the contribution of economic growth to the reduction in early childhood undernutrition in developing countries is very small, if it exists at all." (Vollmer *et al.* 2014, e225)

Abhijeet Singh devotes his comment in the same issue of *The Lancet* to exploring why growth fails so decisively and how public policies geared towards child undernutrition should deal with this apparent reality. Both authors advocate a shift from "the so-called trickle down approach of a growth-mediated strategy" to "direct investments in health and nutrition" (Vollmer *et al.* 2014, e233).

In this note we re-examine the data and show that the association between economic growth and early childhood undernutrition is much stronger than suggested by Vollmer *et al.* (2014). To do so we

**O'Connell: Professor of Economics, Swarthmore College and former Chief Economist, United States Agency for International Development (USAID); Smith: Economist, USAID Bureau for Economic Growth, Education, and Environment.*

address three simple shortcomings in their empirical approach. The first is that the GDP data used in the study are subject to substantial measurement error along precisely the growth dimension that is crucial to their study. In introducing version 8 of the Penn World Tables (PWT8.0), Feenstra *et al.* (2013, 2015) caution against using these data to measure rates of economic growth. Consistent with a measurement-error interpretation, we find that growth coefficients are more than 75 percent higher when using the appropriate national accounts data than when using the PWT8.0 data.

Second, a large proportion of the authors' data comes from surveys that are separated by short periods. Particularly in low- and middle-income countries, short-run changes in GDP are subject to transitory phenomena that include measurement error and cyclical fluctuations. We find that limiting the sample to the longest-available intervals between surveys – an approach that reflects the emphasis of national governments and development partners on growth that is sustained over time – increases estimated growth coefficients further, across all measures of undernutrition.

Third, a very short list of unusual observations plays a powerful role in obscuring the association between growth and childhood undernutrition in the Vollmer *et al.* (2014) study. When we take a conventional approach to identifying such observations, we find that removing them from the sample increases estimated coefficients by nearly 40 percent on average.

To develop these points we focus on the panel of 121 survey-level observations that form the core of the Vollmer *et al.* (2014) study. We do most of the analysis without fixed time effects, and find that when these effects are included they play a key role in reducing the size and statistical significance of the relationship between growth and undernutrition. This is consistent with the presence of unobserved heterogeneity, but also with the distorting impact of measurement error given the very limited temporal sample. The implications for inference are unclear, as we discuss. Very importantly, we find that the robust association between growth and undernutrition re-emerges when we use standard criteria to remove unduly influential outliers.

Our findings are therefore strongly at odds with the conclusions of Vollmer *et al.* (2014). We are also uncomfortable with the underlying thought experiment. As emphasized by Alderman *et al.* (2014), Vollmer *et al.* report the impacts of a 5 percent increase in real GDP per capita. But a single year of 5 percent growth is not what is at stake in debates about economic growth. Sustained growth is what matters, and such growth is transformational, if only via the force of compounding. Over the course of the 2030 Agenda, for example, increasing real GDP per capita by half will require annual growth of only 2.74 percent. Our own results imply that for a country starting with a 50 percent prevalence of stunting, a 15-year episode of growth at this rate would reduce the expected prevalence of stunting by between 6 and 10 percentage points of the relevant age group (using the range of coefficient estimates from Tables 1-3). Growth at 5 percent per year would more than double real GDP per capita over the period and reduce the expected prevalence of stunting by as much as 20 percentage points.[1]

Taken together, these observations favor a far more balanced position than the one embraced by Vollmer *et al.* (2014). References to trickle-down development strategies seem to us particularly misleading. There is little evidence that such strategies have animated Western donor institutions or their private-sector counterparts for any significant part of the past two decades. Growth was not among the Millennium Development Goals, while hunger and health played central roles. More importantly, while economic growth remains an objective of national development plans worldwide and a strategic pillar for some donor institutions, the language of these commitments explicitly repudiates trickle-down. Governments and development partners embrace growth that is sustained, broad-based, and inclusive. This concept of growth has been incorporated into the 2030 Agenda – where, on our interpretation of the evidence, it will make a critically important contribution, alongside direct interventions, to accelerating progress towards early childhood undernutrition goals.[2]

---

[1]See the Appendix and Table 7 in the text.

[2]While the research in this paper was undertaken while both authors were at USAID, the views expressed here are strictly our own (see disclaimer). For an overview of USAID's multi-sectoral nutrition strategy – which draws on the evidence base laid out in Bhutta *et al.* (2013) and Ruel *et al.* (2013), and employs a wide range of instruments including direct health and nutrition

The empirical model

Vollmer *et al.* (2014) model the probability that a young child with given characteristics will be classified as undernourished according to standard age-specific anthropometric thresholds. The sample includes nearly half a million child-level observations drawn from Demographic and Health Surveys (DHS) conducted between 1990 and 2011. But there are only 121 observations on the main explanatory variable of interest, real GDP per capita, because each child in the sample is assigned the national real GDP per capita for the country and year of the survey from which their observation was drawn. The sample includes 36 low-and middle-income countries with at least two DHS surveys between 1990 and 2011. Each country contributes an average of 3.4 surveys to the sample, yielding an average of 5.5 national observations on real GDP per capita per year (Figure 1).

Since we lack the detailed survey data, we will conduct our own analysis at the aggregate level, working with country-level averages for the nutritional outcomes of interest (these are available in the online data appendix to Vollmer *et al.* 2014), rather than with child-level observations. It bears emphasis that despite the compression of information on nutritional outcomes that is implied by going from individual-level to aggregate data, the economic growth information we bring to bear on identifying the nutrition impacts has precisely the same domain – varying only across surveys and not across children – as that of the authors.

Using p_{it} to denote the overall prevalence of stunting, wasting, or underweight in country i and survey year t, therefore, we will estimate logistic regression models of the form

$$ln\frac{p_{it}}{1-p_{it}} = \gamma \ln y_{it} + \mu_i + \delta_t + \varepsilon_{it},$$

where y_{it} is country-wide real GDP per capita and μ_i and δ_t are unobserved country and survey-year constants. For reasons explained below, we augment the country-level dataset by bringing in real GDP per capita at constant local currency from the World Bank's *World Development Indicators*.

programming along with interventions to empower women and promote inclusive economic growth in agriculture – see USAID (2014).

The dependent variable in these regressions is the log of the odds of being stunted, wasted, or underweight.[3] Along with the log of real GDP per capita, Vollmer *et al.* (2014) include a variety of additional covariates drawn from the survey data, including the child's birth order, urban location, mother's education, and others. Lacking the survey data, we omit these and focus on the bivariate relationship between GDP per capita and childhood undernutrition. We estimate by ordinary least squares (OLS) and report significance tests using standard errors clustered at the country level.[4] Our interest is in γ, the coefficient on the log of real GDP per capita.

Crucially, we follow Vollmer *et al.* (2014) in including a full set of country dummy variables. These allow us to control for any time-invariant *country fixed effects* (μ_i) that influence the odds but may be correlated with real GDP per capita. Failing to control for this type of unobserved heterogeneity would bias the estimate of γ.

Computationally, the inclusion of country fixed effects converts a regression that is specified and estimated in levels into a numerically equivalent regression that is specified in *within-country differences* of the variables, defined as the deviations of each the variables from their within-country in-sample means. This has two important implications. The first is that it is no longer important to measure the levels of real GDP on a comparable basis across countries, because the country-level averages of all variables have been swept out of the data. The sample information is limited to within-country differences in real income per capita over time, and therefore to the effects of cumulative economic growth or decline, rather than the effects of cross-

[3] Stunting, wasting and underweight correspond to height for age, weight for height, and weight or age more than 2 standard deviations below a reference global median. Height for age is sometimes viewed as an indicator of chronic undernutrition, weight for age acute undernutrition, and weight for height either chronic or acute or both.

[4] With the exception of F tests for joint significance of time dummies, the qualitative results in terms of statistical significance do not depend on whether standard errors are clustered at the country level or estimated using Stata 13's 'robust' option. The F tests are uniformly stronger in rejecting the null of no joint impact of time effects when standard errors are clustered at the country level.

country differences in living standards.[5] This means that the coefficient on real GDP per capita is validly interpreted as a growth effect – and that what matters for getting an accurate estimate of this coefficient is getting growth rates right, not getting average levels right. The difference is important, as we will see.

The second implication is directly related. The inclusion of country fixed effects protects against bias coming from unobserved heterogeneity by throwing away the cross-country or *between-groups* variation in the data. The cost of this loss of information is that any measurement error in the level of real GDP is exacerbated, because it translates into much greater proportional error in growth rates. The impact of this increased measurement error is to bias the fixed-effects estimate of γ towards zero (Griliches and Hausman 1986). This effect increases the premium on avoiding measurement errors in growth rates.

We develop our first three points in regressions that exclude time effects. Table 5 then follows Vollmer *et al.* (2014) in also including a full set of yearly dummy variables.

Addressing measurement error

The authors observe in passing that real income per capita in Nigeria grew at the rate of 18.7 percent per year between the DHS surveys conducted 2003 and 2008. This astonishing observation would imply an increase in Nigerian GDP per capita of 136 percent in 5 years.[6] A glance at the online Appendix reveals even bigger puzzles – for example, Nigeria is recorded as having grown at 30 percent per year between 1999 and 2003, implying that Nigeria's GDP per capita more than tripled in 4 year period. In Bangladesh, real GDP per capita falls between 1996 and 2007 in the authors' data, a period during which the national accounts record cumulative growth of over 50 percent.

These growth puzzles are an artifact of a major methodological change that was built into the new PWT8.0 data and is documented in detail by Feenstra *et al.* (2013, 2015). In the past, each re-benchmarking of the Penn World Tables (at roughly 5-year intervals) has provided a newly-accurate snapshot of relative living standards in the benchmark year, and has then been extrapolated backwards and forwards – thereby replacing the previous PPP-adjusted time series – using the growth rates of real GDP per capita from each country's national accounts. By construction, therefore, in any given year the time-series data on PPP-adjusted real GDP per capita at constant international prices (as downloaded, for example, from the *World Development Indicators*) have displayed the same inter-annual growth rates as the time-series data on real GDP per capita in constant local currency.

As observed by Johnson *et al.* (2013), one result of this procedure is that each re-benchmarking of the Penn World Tables creates the appearance of invalidating the relative living standards implied by earlier benchmarks. In a major departure from previous practice, PWT 8.0 neutralizes this effect by 'stitching together' successive benchmark years. In doing so, however, it produces an interpolated set of cross-country snapshots, rather than a single snapshot augmented with country-specific national accounts growth rates. The authors of the new Penn World Tables are very clear about what this new approach implies for researchers:

> "The new method of estimating PPPs has arguably led to a measure of real GDP that is more reliable than before since older benchmark information is no longer discarded. This has substantially changed PWT data, as benchmark and interpolated observations now cover one-third of all observations in PWT. As a consequence, though, [PPP-adjusted] real GDP has become less suitable to measure changes over time in a single country. Real GDP has always been less than ideal for this purpose, as it is estimated using information on spending patterns across all countries. Since a country's spending pattern is a result of its own preferences and relative prices, other countries' spending patterns are irrelevant when measuring the economic performance of a single country over time. So if an

[5] The growth interpretation is further underscored by the fact that if intra-survey intervals were constant across countries and periods, these regressions would be (virtually) numerically equivalent to regressions specified in intra-survey time differences of the variables.

[6] The authors compound this puzzle by referring to the Nigerian case in their response to Alderman *et al.* (2014). See Vollmer *et al.* (2014b).

> analysis aims to explain cross-country differences in GDP *growth rates*, we would strongly recommend using data on the growth of GDP at constant national prices, based directly on a country's National Accounts." (Feenstra *et al.* 2015, pp. 23-24; italics in original)

The distinction between growth rates as a dependent variable (as in the quotation) and growth rates as an explanatory variable (as in Vollmer *et al.* 2014) is immaterial: when growth rates of real GDP per capita are the object of study, the PWT 8.0 data are inappropriate.[7]

As suggested by the Nigerian data, the differences between the PWT8.0 growth rates and the national accounts growth rates are large. For the country/year observations in the Vollmer *et al.* (2014) sample, the correlation between the intra-survey growth rates generated by the two series is only 0.52. This is far below the already-low global correlation of 0.71 between inter-annual growth rates in the two series (Feenstra *et al.* 2013a): the difference is probably a reflection of the poorer fit of the international reference basket to the economic structures of low- and middle-income countries, as well as the lower quality of the GDP data in these countries (Feenstra *et al.* 2013b). As indicated in Figure 2, a regression of the constant-local-currency growth rates on the PWT8.0 growth rates yields a coefficient that is not significantly different from 1. The PWT8.0 data therefore provide an unbiased estimate of constant-local-currency growth rates, but one that contains a very large measurement error.

The second segment of Table 1 shows the impact of replacing the PWT8.0 growth rates with the national accounts growth rates. The national accounts figures would of course be inappropriate in a pure levels regression, because constant-local-currency data are index numbers that are not comparable across countries. As we have emphasized, however, the country dummy variables sweep out all between-country comparisons from the data. Only the growth information is retained. Columns 4-6 are therefore equivalent to the authors' regressions, only with less measurement error in the variable of interest. All three coefficients rise very substantially – by between 64 and 93 percent, or 79 percent on average – and inference becomes uniformly more precise.

Focusing on longer episodes

The national accounts data are also, of course, subject to measurement errors, particularly among low- and middle-income countries. If these are partly transitory, the estimated impact of economic growth will be biased towards zero to a degree that depends on the weight of short-interval comparisons in the sample. A simple way to assess this effect is to restrict attention to the longest available inter-survey interval for each country. We do this in Table 2, running regressions that retain only the first and last survey in the sample for each country. By comparison with the first panel of Table 1 (the Lancet specification), each of the coefficients more than doubles in size, for an average increase of 125 percent. There is little or no reduction in the statistical precision of the estimates when we focus on the longest spells, despite the much smaller sample size.

While the results in Table 2 are consistent with a measurement-error interpretation, they are also consistent with the possibility that the effects of economic growth on undernutrition are smaller per 'point-year' of growth when growth is transitory than when it is sustained. There are good reasons to suspect that this type of duration dependence may be present in the data, given the impact of cumulative growth or decline on asset buffers and other mechanisms for stabilizing the household's living standards. The statistical impact of transitory measurement error is also duration-dependent, however, so disentangling these two interpretations of duration dependence is a nontrivial task and we do not attempt it here.

We caution that the results in Table 2 do not imply that transitory macroeconomic shocks are of limited concern from the perspective of child nutrition. It is well documented that recessions tend to be more severe and protracted in low- and middle-income countries than in industrial countries (Hausmann *et*

[7] Feenstra, Inklaar and Timmer are more explicit in their VOX-EU post of 2 September 2013: "These results imply, in our view, that real GDP measures should be used not as a substitute for National Accounts growth figures for a country, but instead as measures of relative income or output across countries." (http://www.voxeu.org/article/recasting-international-income-differences-next-generation-penn-world-table)

al. 2008). If large transitory shocks produce medium-term movements in GDP, then – as in the case of measurement error – the information content of medium-term comparisons is likely to be stronger than that of short-term ones in determining the impact of growth on undernutrition.

Managing outliers

Figure 3 shows a scatter plot of the within-groups information – the deviations of the variables from their country means – on the association between underweight and growth. The within-groups variation is the only information being brought to bear in a regression that includes country-level dummy variables; the OLS fit through this scatter plot yields exactly the coefficient reported in the first panel of Table 2. The downward slope captured by the regression is clearly evident in the data, but so is the relatively wide spread of the observations and the resulting imprecision of the estimate.

Visually, several of the observations in the scatter plot stand out as potentially influential outliers: Armenia's 2000 observation is a dramatic case in point. To identify unusually influential observations in a more systematic fashion, we report in the Appendix two measures of the *influence* of individual observations on the estimated OLS coefficient (Cook's distance and DFBETA). For completeness we also report two measures of the *discrepancy* between individual data points and the estimated regression line (Studentized and standardized residuals). We calculate these measures for all three undernutrition variables. We then re-estimate each regression twice: first, excluding the observations that exceed conventional hurdles for the two influence measures for that particular dependent variable; and second, excluding the set of *all* observations that exceed these influence hurdles for at least one of the three dependent variables. Table 3 in the Appendix provides details, and lists the excluded observations.

There is no generally accepted standard for whether observations with unusually high influence should be retained or removed from a regression sample. What is clear from Table 4, however, is that a small number of unusual observations are playing an extremely powerful role in reducing the estimated impact of growth on child undernutrition. Eliminating these observations generates a major increase in coefficient sizes – more than a *tripling* on average (320 percent), by comparison with their results – along with a large gain in statistical precision. The wasting coefficient is now statistically significant along with the others, for the first time in these results.

Time effects matter

Table 5 augments the specification in the first panel of Table 2 to include a full set of yearly time dummies. The rationale for including time dummies is to eliminate spurious correlations that may be generated by unobserved temporal heterogeneity – i.e., by variables that are constant across countries but that vary (are heterogeneous) over time. The difficulty this creates is that as long as these variables are unobserved, there no way to eliminate their influence without throwing away a key part of the temporal variation in the data. The sample information on growth is already restricted to 121 observations on intra-survey growth rates, so this is potentially very costly in terms of detecting the statistical relationship of interest. Fixed-effects regressions are subject to substantial attenuation of coefficients in the presence of measurement error, because the elimination of between-groups variation reduces the ratio of signal to noise in the data (Griliches and Hausman 1986). Figure 4 shows how sharply the variation in the data falls as within-groups and then within-periods variation is eliminated.

Not all unobserved temporal heterogeneity, of course, is damaging. To generate bias, an unobserved variable must matter directly for country-level undernutrition, meaning that its effect must not be solely mediated through growth; and it must be correlated with country-level growth. One example might be an international child-health campaign lasting several years that was large enough to influence undernutrition and was positively correlated (even by happenstance, given the short sample) with global economic growth. Another might be global technological advances that simultaneously reduced undernutrition and generated global economic growth.

With time effects omitted, global variables like health campaigns and technological advances would implicitly be in the residual, where a positive correlation with country-level growth would produce an over-estimate of the growth coefficient. The danger, of course, is that in throwing out *all* such variation, we are also eliminating information that may be crucial for identifying the impact of growth in small samples. For example, the global business cycle has a strong impact on growth in low- and middle-income countries, but probably has very little direct effect on childhood undernutrition. This component of sample-wide growth would not be not needed if the sample were long, but it may be crucial to identifying the impact of country-level growth when there are only a few temporal observations per country.

In Table 5, the inclusion of time effects reduces estimated effect sizes and weakens the statistical significance of the coefficients. The stunting result remains robust, both in terms of size and in statistical significance when using the national accounts data. But the wasting and underweight coefficients shrink dramatically in size and the latter is now statistically insignificant (the wasting coefficient was already statistically insignificant in column 5 of Table 1).

The inclusion of unobserved time effects therefore appears to be crucial to the weak statistical results in the Vollmer *et al.* (2014) study. As indicated in Table 6, however, this effect is remarkably sensitive to the influence of outlier observations. As before, we do not choose the outliers that are excluded in Table 6 selectively: we identify them using a standard and automatic screen based on undue influence (see Table 4). When these observations are omitted, the estimated coefficients are easily as large as they were when time effects were excluded. The coefficients are also strongly statistically significant – a result that holds for all three measures of child undernutrition. Figure 5 provides a clear visual illustration of the impact of a few unduly influential observations.

Conclusions

Vollmer *et al.* (2014) present their paper as evidence that growth is irrelevant for reducing childhood undernutrition. They conclude that progress will require a nearly-exclusive focus on public health programs that directly target childhood undernutrition. We do not think the evidence bears this interpretation. We readily concede that the relationship between national income and measures of health and morbidity is not automatic (Ruel *et al.* 2013); and there is certainly no basis in the global evidence to argue that advances in health and nutrition in a given country are fully dependent on whether economic growth occurs in that country (Deaton 2013). But it is not true that economic growth has virtually no association with early childhood undernutrition among low- and middle-income countries. We find the opposite. There is an appreciable association even within the authors' own sample, when we use an appropriate measure of real GDP. The results are stronger yet when we eliminate short-term variation and focus on sustained growth or decline. Finally, a few unusual observations play an extremely important role in reducing estimated coefficients; with these observations excluded, the relationship is strong and statistically significant.

Table 7 reassesses the "quantitatively very small to null association" reported in the Lancet study against the evidence reported in this note. What emerges clearly here is that the conclusions of Vollmer *et al.* (2014) reflect a combination of downward bias in the estimated coefficients and limited scope in the underlying thought experiment. With a coefficient of -0.2 and a mere 5 percentage points of added GDP, the reduction in undernutrition is indeed small. But larger coefficients generate proportional increases in the associated decline in undernutrition, and of course a great deal depends on the duration of growth differences. Using our own coefficient estimates and starting at an incidence rate of 50 percent, five percent real income growth per capita over the remainder of the 2030 Agenda – which would more than double real GDP per capita over the course of 15 years – would be associated with a reduction in undernutrition equivalent to between 10 and 20 percent of the population of children in low- and lower-middle-income countries. Even 1 percent growth, over that period, would be associated with three times the cumulative decline in undernutrition that a single year of 5 percent growth would accomplish.

Time effects weaken these relationships in the full sample, a finding that points to some combination of measurement error and unobserved heterogeneity. In the case of stunting, however, they do not eliminate it; and, as we have argued, the inclusion of time effects throws away variation that may be crucial to identifying the empirical relationship in small samples. There is a well-known tradeoff here between the downward bias of fixed effects under measurement error, and the upward bias from unobserved time-based heterogeneity. In our view, when the temporal dimension of the panel is as limited as it is in the present study, some weight should be given to results that exclude the time effects. More importantly in the present context, we find that there is a powerful interaction between the time effects and the influence of a few egregious outliers. Eliminating these outliers delivers a robust association between early childhood undernutrition and economic growth even when the time effects are included.

Acknowledgments

Stephen O'Connell is Gil and Frank Mustin Professor of Economics at Swarthmore College, and was Chief Economist of USAID when this research was initiated. Caroline Smith is an Economist in the Economic Policy Office of USAID. The authors are grateful to Robert Wuertz and Betsy Jordan-Bell for helpful discussions and to Colin Huerter for expert research assistance. Comments are welcome.
soconne1@swarthmore.edu carsmith@usaid.gov

Appendix: Data and Methods

Logistic regressions model the probability that a dichotomous dependent variable that takes on values of either 0 or 1 is equal to 1, conditional on a set of explanatory variables. For example, the dependent variable may be an indicator for whether the *i*th child in country *j* and survey year *t* is stunted $(y_{ijt} = 1)$ or not$(y_{ijt} = 0)$. The expected value of y_{ijt} conditional on the vector $\boldsymbol{x}_{ijt}$, which is equal to the conditional probability that $y_{ijt} = 1$, is modeled as the logistic function

$$E[y_{ijt}|x_{ijt}] = p_{ijt} = \frac{\exp(\boldsymbol{x}_{ijt}'\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_{ijt}'\boldsymbol{\beta})},$$

where $p_{ijt} = Prob[y_{ijt} = 1 \mid x_{ijt}]$. Alternatively, logistic regressions model the odds that $y_{ijt} = 1$ as a log-linear function of the $\boldsymbol{x}$ vector. The above equation implies

$$odds_{ijt} \equiv \frac{p_{ijt}}{1-p_{ijt}} = \exp(\boldsymbol{x}_{ijt}'\boldsymbol{\beta}).$$

When applied to micro data, the parameter vector $\boldsymbol{\beta}$ can be estimated by maximum-likelihood methods. If the dependent variables are already measured as aggregate frequencies, however – for example, as survey-wide rates of stunting for individual countries and years – then the odds can be calculated for each country and year in the data, and the coefficients can be recovered from the OLS regression

$$\ln(odds_{jt}) = \boldsymbol{x}_{jt}'\boldsymbol{\beta}.$$

Vollmer *et al.* (2014) use survey-level micro data on individuals and therefor estimate their models by maximum likelihood. We use country-level survey means for the same countries and years, and apply OLS. In either case, the estimated coefficients from this regression give the (approximate) percentage change in the odds ratio from a one-unit change in the *k*th explanatory variable.

In their simplest ("unadjusted") specification, Vollmer *et al.* (2014) allow for country and time dummy variables ($C(j)$ and $T(t)$, respectively) and for a single time-varying country-level variable, the log of real GDP per capita.[8] The set of explanatory variables is therefore

$$\boldsymbol{x}_{jt}'\boldsymbol{\beta} = \mu + \sum_{j=1}^{N-1} \mu_j \cdot C(j)_{jt} + \sum_{t=1}^{T-1} \delta_t \cdot T(t)_{jt} + \gamma \ln y_{jt},$$

where we have omitted one country and one time period from the list of dummy variables in order to avoid perfect collinearity with the constant.

The coefficient γ is the elasticity of the odds ratio with respect to real GDP per capita. To translate this into the impact of real GDP on the probability p_{jt} we can use

$$\begin{aligned} dp_{jt}/d\ln y_{jt} &= \frac{dp_{jt}}{d\ln odds_{jt}} \cdot \frac{d\ln odds_{jt}}{d\ln y_{jt}} \\ &= \frac{dp_{jt}}{d\ln odds_{jt}} \cdot \gamma \\ &= [p_{jt}(1-p_{jt})] \cdot \gamma. \end{aligned}$$

The impact therefore depends on the value of p_{jt} and is largest when $p_{jt} = 1/2$.

Vollmer *et al.* (2014) estimate their model using a Stata 13 command that reports the estimated *odds ratios* for $\ln y_{jt}$ rather than the estimated coefficient. In the general case of $M \geq 1$ explanatory variables including the constant term, the odds ratio for the kth variable is defined as the ratio of the odds when $x_k = x_{k0} + 1$ to the odds when $x_k = x_{k0}$:

[8] The 'adjusted' results include additional micro-level explanatory variables, like whether the child is in an urban location.

$$or_{ijt}(1\ extra\ unit\ of\ x_k) =$$

$$\frac{\left[\prod_{m\neq k}\exp\left(b_m x(m)_{ijt}\right)\right]\cdot \exp(b_k\cdot[x_{k0}+1])}{\left[\prod_{m\neq k}\exp\left(b_m x(m)_{ijt}\right)\right]\cdot \exp(b_k\cdot x_{k0})}$$

$$= \exp(b_k).$$

This has standard error $\exp(b_k)\cdot s_k$, where s_k is the estimated standard error of b_k.

This calculation can be generalized to give the odds ratio for any given change in an explanatory variable. For example, Vollmer *et al.* (2014) report the odds ratios for a 5 percent increase in real GDP per capita. After cancelling out the other explanatory variables (as in the above expression), this is given by

$$or_{ijt}(5\%\ \text{increase}) = \frac{\exp(\gamma\cdot\ln(1.05y_{ijt}))}{\exp(\gamma\cdot\ln(y_{ijt}))} = \frac{\exp(\gamma\ln(y_{ijt})+\gamma\ln(1.05y_{ijt}/y_{ijt}))}{\exp(\gamma\ln(y_{ijt}))} = \exp[\gamma\ln(1.05)].$$

This gives us two equivalent expressions for the odds ratio associated with a 5 percent increase in income:

$$or_{ijt}(5\%\ \text{increase}) = [\exp(\gamma)]^{\ln(1.05)} = 1.05^{\gamma}.$$

For any specified growth factor G, these odds ratios take the form

$$or_{ijt}(G-1\ percent\ increase) = [\exp(\gamma)]^{\ln(G)} = G^{\gamma}.$$

These quantities are independent of the starting level of income. They also compound over time: if 5 percent annual growth is maintained for T years, then the corresponding growth factor is 1.05^T and the odds ratio is $1.05^{T\gamma}$. The exponent to apply to the growth factor is the coefficient γ multiplied by the number of years of growth.

To calculate the standard error of the odds ratio, there are a variety of approaches to getting the standard error of a nonlinear function. One practical approach is to 'trick' Stata into delivering the standard error directly, by running the model with the log of GDP per capita replaced by $20\cdot\ln y_{ijt}$. A one-unit increase in this new variable is exactly equal to a 5 percent increase in real GDP per capita. Alternatively, in the simple case of aggregate data, where we are just doing a linear regression of the log of the odds, the estimated coefficient on the log of real GDP per capita would simply have to be divided by 20, and similarly for its estimated standard error (leaving the t-statistic and inference unchanged).

Identifying unusual observations

Stata includes a number of post-estimation commands to help identify potentially troublesome observations. Standardized and Studentized residuals are methods of adjusting residuals for their standard errors. These measures are indicators of *discrepancy*: they seek to characterize the degree to which a particular observation is in line with the other observations. For further discussion of standardized and Studentized residuals, see Chatterjee and Hadi (1988).[9] In identifying unusual observations, we employ a cutoffs of 3 for the standardized and Studentized residuals.

Cook's distance and DFBETA, on the other hand, are measures of *influence*. Influence takes into consideration both *leverage* (a measure of how extreme an observation is in the space of the explanatory variables) and discrepancy. Influence shows the extent to which regression coefficients change when individual observations are dropped from the sample. Full explanations of Cook's distance and DFBETA can be found in Cook (1977) and Belsley, Kuh, and Welch (1980), respectively. We employ influence cutoffs of $2/\sqrt{n}$ for DFBETA and $4/n$ for Cook's distance, where n is the sample size.

[9] While Stata uses the terms standardized and Studentized residuals, Chatterjee and Hadi (1988) use "internally Studentized" and "externally Studentized" residuals to describe these measures, respectively.

References

Alderman, Harold, Lawrence Haddad, Derek D. Headley and Lisa Smith (2014) "Association between Economic Growth and Early Childhood Nutrition" *The Lancet Global Health* 2(9), e500

Belsley, David A., Edwin Kuh and Roy E. Welch (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity* (New York: Wiley)

Bhutta, Zulfiqar A., Jai K. Das, Arjumand Rizvi, Michelle F. Gaffey, Neff Walker, Susan Horton, Patrick Webb, Anna Lartey, Robert E Black, The Lancet Nutrition Interventions Review Group, and the Maternal and Child Nutrition Study Group (2013) "Evidence-Based Interventions for Improvement of Maternal and Child Nutrition: What Can Be Done and At What Cost?" *The Lancet* 382(9890), August 3: 452–477

Chatterjee, Samprit and Ali S. Hadi (1988) *Sensitivity Analysis in Linear Regression* (New York: Wiley)

Cook, R. Dennis (1977) "Detection of Influential Observations in Linear Regression *Technometrics* 19: 15-18

Deaton, Angus (2013) *The Great Escape: Health, Wealth, and the Origins of Inequality* (Princeton University Press)

Feenstra, Robert C., Robert Inklaar and Marcel Timmer (2015) "PWT 8.0 – A User's Guide" University of California, Davis and University of Groningen

Feenstra, Robert C., Robert Inklaar and Marcel Timmer (2013) "The Next Generation of the Penn World Table" *NBER Working Paper* No. 19255

Griliches, Zvi and Jerry A. (1986) "Errors in Variables in Panel Data" *Journal of Econometrics* 31: 930-118

Hausmann, Ricardo, F. Rodriguez and R. Wagner (2008) "Growth Collapses" in Carmen M. Reinhart, Andres Velasco and Carlos Vegh, eds, *Money, Crises and Transition: Essays in Honor of Guillermo* Calvo (Cambridge: MIT Press)

Johnson, Kiersten (2007) "Migration, Economy and Policy: Recent Changes in Armenia's Demographic and Health Indicators: Further Analysis of Demographic and Health Surveys Data" Macro International Inc. Calverton, Maryland, USA

Johnson, Simon, William Larson, Chris Papageorgiou and Arvind Subramanian (2013) "Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates," *Journal of Monetary Economics* 60(2): 255-274.

Richardson, Erica (2013) "Armenia: Health System Review" *Health Systems in Transition* 15(4), European Observatory on Health Systems and Policies

Ruel, Marie T., Harold Alderman, and The Maternal and Child Nutrition Study Group (2013) "Nutrition-Sensitive Interventions and Programmes: How Can They Help to Accelerate Progress in Improving Maternal and Child Nutrition?" *The Lancet* 382(9891), August 10: 536–551

Singh, Abhijeet (2014) "Why are Economic Growth and Reductions in Child Undernutrition so Weakly Correlated—and What Can Public Policy Do?" *The Lancet Global Health* 2(4), April: e185 - e186

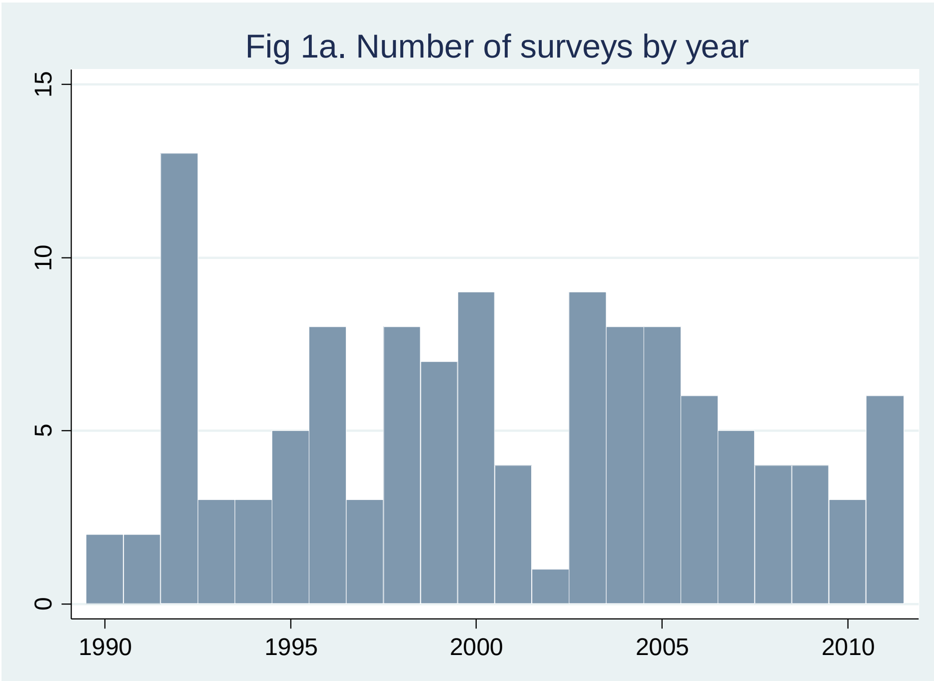
Vollmer, Sebastian, Kenneth Harttgen, Malavika A. Subramanyam, Jocelyn Finlay, Stephan Klasen, and S. V. Subramanian (2014) "Association Between Economic Growth and Early Childhood Undernutrition: Evidence from 121 Demographic and Health Surveys from 36 Low-Income and Middle-Income Countries" *The Lancet Global Health* 2(4), April: e225 - e234

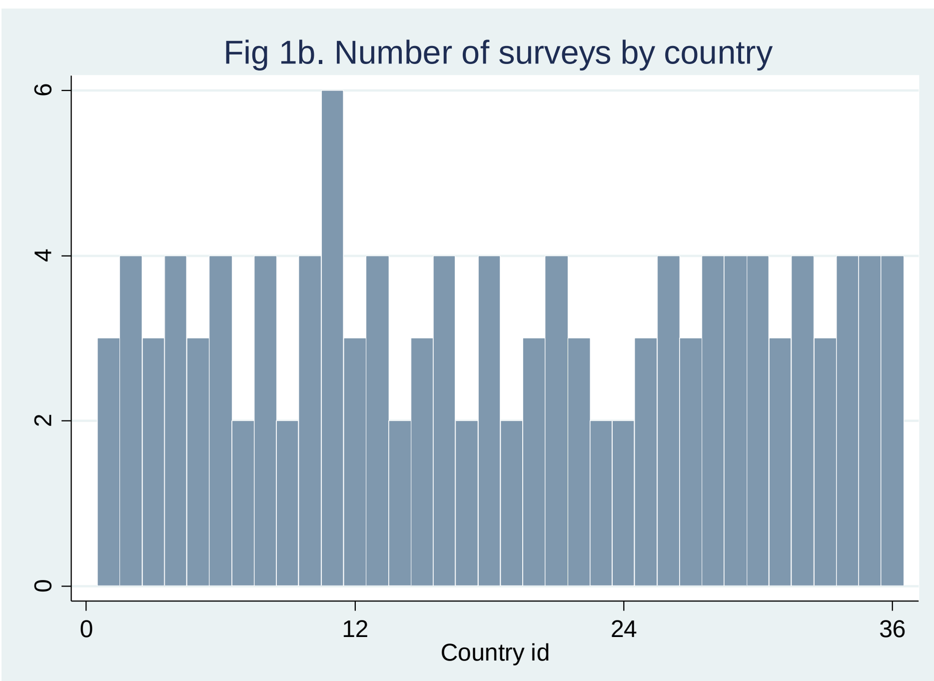
Vollmer, Sebastian, Kenneth Harttgen, Malavika A. Subramanyam, Jocelyn Finlay, Stephan Klasen, and S. V. Subramanian (2014b) "Association Between Economic Growth and Child Nutrition – Authors' Reply" *The Lancet Global Health* 2(9), September: e501-e502

USAID (2014) *Multi-Sectoral Nutrition Strategy, 2014-2025*

Tables and Figures

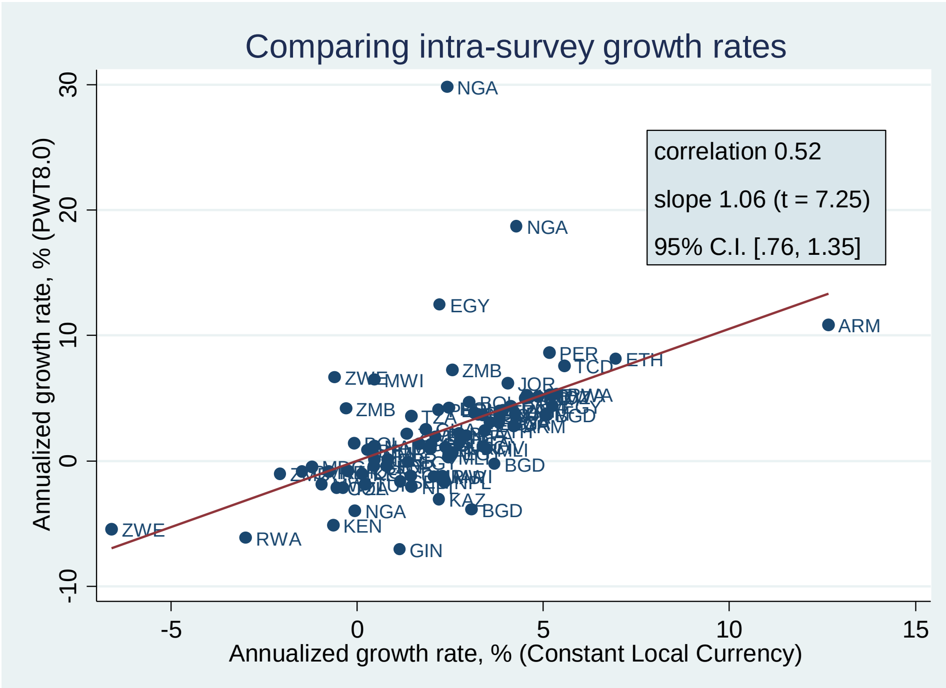
Figure 1 *Distribution of national surveys over time and across countries*





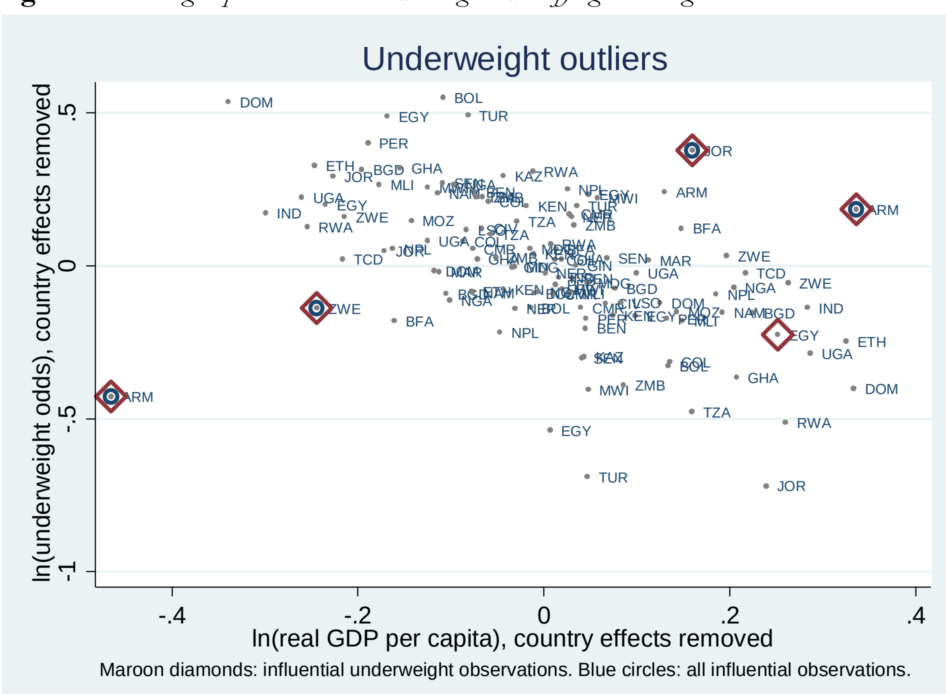
Source: The country aggregates for each survey are reported in the online appendix to Vollmer *et al.* (2014).

Figure 2 *Classical measurement error in GDP growth*



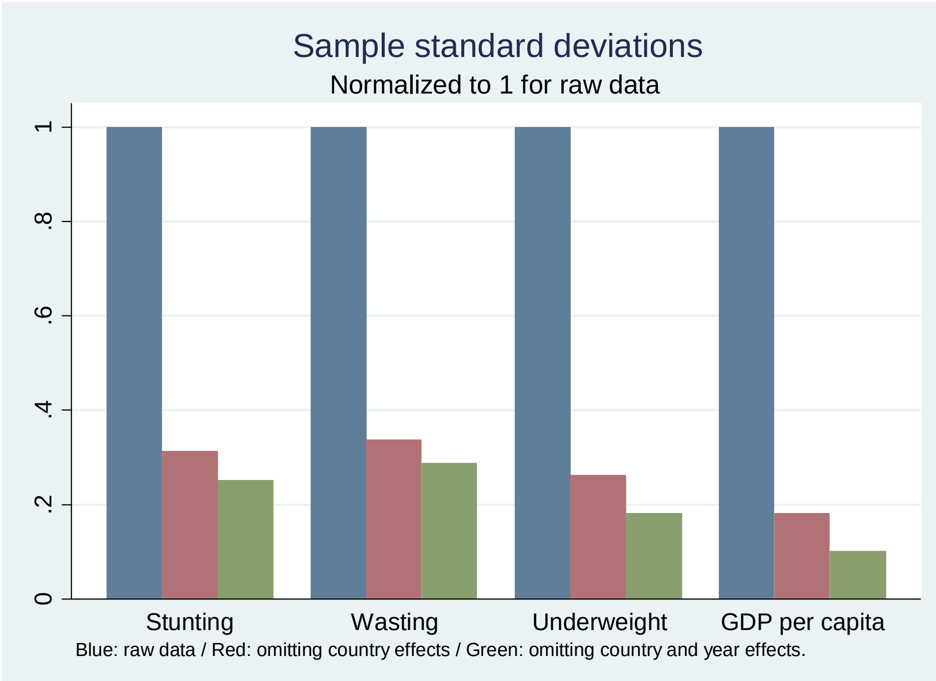
SourceS: PWT8.0 is from version 8 of the Penn World Tables. Constant-local-currency GDP per capita is from the World Bank, *World Development Indicators* online.

Figure 3 *Within-groups variation in underweight: identifying troubling observations*



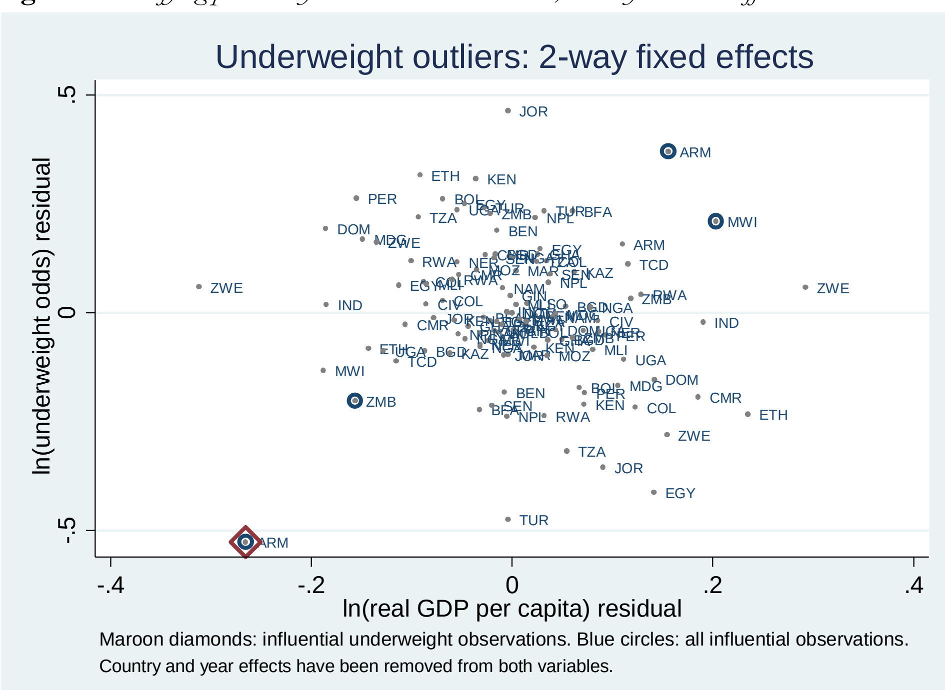
Notes: Observations revealed to be problematic in the underweight regression are marked with a blue circle. Outliers revealed to be problematic in at least one of the three undernutrition regressions are marked with a red diamond.

Figure 4 *Eliminating information by conditioning on countries and periods*



Notes: The figure shows the standard deviations of (i) the variable itself (logs of the odds ratios and of real GDP per capita), (ii) the residual from a regression on country effects, and (iii) the residual from a regression on country and year effects, each as a ratio to the standard deviation of the variable. For real GDP per capita we use PPP-adjusted data (from *World Development Indicators* and based on PWT version 7.x, rather than 8.0) so that the raw standard deviation is meaningful. As explained in the text, this earlier version of the PPP-adjusted data displays the same growth rates as the constant-local-currency data.

Figure 5 *Identifying potentially troublesome observations, country and time effects included*



Notes: See note to Figure 3.

Table 1 *Addressing classical measurement error*

Variable	Stunting	Wasting	Under-weight	Stunting	Wasting	Under-weight
$\ln y_{it}$ (PWT8.0)	**-0.338*****	**-0.148**	**-0.395***			
	-3.296	-0.849	-1.874			
	0.002	0.402	0.069			
$\ln y_{it}$ (Nat'l Accts)				**-0.605*****	**-0.243**	**-0.761****
				-3.463	-0.992	-2.622
				0.001	0.328	0.013
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Time effects	No	No	No	No	No	No
N	121	121	121	121	121	121
R^2	0.916	0.888	0.939	0.926	0.888	0.946
Adjusted R^2	0.880	0.840	0.913	0.894	0.840	0.923

Notes: t-statistics are below coefficients, and significance levels are below those. Standard errors are clustered at the country level. Asterisks (*, **, ***) indicate significance at the 10%, 5% and 1% levels, respectively.

Table 2 *Focusing on the longest intra-survey spell in each country*

Variable	Stunting	Wasting	Under-weight
$\ln y_{it}$ (Nat'l Accts)	**-0.706****	**-0.362**	**-0.875****
	-2.629	-1.168	-2.139
	0.013	0.251	0.040
Country effects	Yes	Yes	Yes
Time effects	No	No	No
N	72	72	72
R^2	0.941	0.901	0.946
Adjusted R^2	0.881	0.800	0.891

Notes: t-statistics are below coefficients, and significance levels are below those. Standard errors are clustered at the country level. Asterisks (*, **, ***) indicate significance at the 10%, 5% and 1% levels, respectively.

Table 3 *Identifying potentially troublesome observations*

Observation	Stunting				Wasting				Underweight			
	1-way		2-way		1-way		2-way		1-way		2-way	
	I	D	I	D	I	D	I	D	I	D	I	D
Armenia 2000	x		x		x		x		x	X	x	x
Armenia 2010	x								x		x	
Egypt 2008	x				x							
Jordan 1997	x											
Jordan 2007					x	x			x			
Madagascar 2004			x									
Malawi 2000											x	
Rwanda 2000	x											
Zambia 2007											x	
Zimbabwe 1999					x		x					
Zimbabwe 2011					x		x		x			
Fixed effects												
Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes

Notes: 1-way and 2-way refer to regressions with fixed country effects only, or with both country and year effects (as indicated). I denotes unusual influence (DFBETA > $2/\sqrt{n}$ and Cook's Distance > $4/n$), and D denotes unusual discrepancy (standardized and Studentized residuals > 3).

Table 4 *Removing potentially troublesome observations*

Variable	Regression-specific outliers			All outliers		
	Stunting	Wasting	Under-weight	Stunting	Wasting	Under-weight
$\ln y_{it}$ (Nat'l Accts)	**-0.858*****	**-0.643****	**-1.106*****	**-0.927*****	**-0.576****	**-1.150*****
	-5.723	-2.542	-6.905	-6.168	-2.403	-6.322
	0.000	0.016	0.000	0.000	0.022	0.000
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Time effects	No	No	No	No	No	No
N	116	116	117	113	113	113
R^2	0.939	0.918	0.960	0.941	0.919	0.959
Adjusted R^2	0.911	0.880	0.942	0.913	0.880	0.939

Notes: t-statistics are below coefficients, and significance levels are below those. Standard errors are clustered at the country level. Asterisks (*, **, ***) indicate significance at the 10%, 5% and 1% levels, respectively. Regression-specific outliers are identified in the "1-way" columns of Table 3. The "All outliers" columns eliminate the full set of 1-way outliers in Table 3, regardless of the regression in which they appear.

Table 5 *Incorporating time effects*

Variable	Stunting	Wasting	Under-weight	Stunting	Wasting	Under-weight
$\ln y_{it}$ (PWT8.0)	**-0.203** -1.611 0.116	**-0.222** -1.002 0.323	**-0.148** -0.851 0.401			
$\ln y_{it}$ (Nat'l Accts)				**-0.424*** -1.833 0.075	**-0.023** -0.060 0.952	**-0.221** -0.585 0.562
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F(21,35) for time effects	0.0000	0.0133	0.0000	0.0495	0.0029	0.0000
N	121	121	121	121	121	121
R^2	0.940	0.918	0.968	0.942	0.917	0.968
Adjusted R^2	0.885	0.845	0.938	0.889	0.842	0.938

Notes: t-statistics are below coefficients, and significance levels are below those. Standard errors are clustered at the country level. Asterisks (*, **, ***) indicate significance at the 10%, 5% and 1% levels, respectively.

Table 6 *Incorporating time effects and removing potentially troublesome observations*

Variable	**Regression-specific outliers**			**All outliers**		
	Stunting	Wasting	Under-weight	Stunting	Wasting	Under-weight
$\ln y_{it}$ (Nat'l Accts)	**-0.643***** -2.792 0.008	**-0.730*** -1.827 0.076	**-0.683***** -2.795 0.008	**-0.896***** -3.254 0.003	**-0.711*** -1.726 0.093	**-0.911***** -3.067 0.004
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
Prob >F(21,35) for time effects	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
N	119	118	117	114	114	114
R^2	0.947	0.929	0.976	0.949	0.934	0.978
Adjusted R^2	0.897	0.862	0.954	0.896	0.867	0.955

Notes: t-statistics are below coefficients, and significance levels are below those. Standard errors are clustered at the country level. Asterisks (*, **, ***) indicate significance at the 10%, 5% and 1% levels, respectively.
Regression-specific outliers are identified in the "2-way" columns of Table 3. The "All outliers" columns eliminate the full set of 2-way outliers in Table 3, regardless of the regression in which they appear. Note that a country must have at least 2 observations to contribute information to the regression; so if Armenia's 2000 and 2010 observations are omitted the 2005 observation is also effectively omitted (by being fit perfectly by the country dummy).

Table 7 *Reduction (percentage points) in the incidence of undernutrition starting at 50%*

Coefficient	Cumulative growth factor			
	1.05	1.16	1.50	2.00
-0.20	0.2	0.7	2.0	3.5
-0.40	0.5	1.5	4.0	6.9
-0.60	0.7	2.2	6.1	10.2
-0.80	1.0	3.0	8.0	13.5
-1.00	1.2	3.7	10.0	16.7
-1.10	1.3	4.1	11.0	18.2
	Implied constant annual growth rate (%)			
if over 1 year	5	16	50	100
if over 15 years	0.3	1.0	2.7	4.7

Notes: The table shows the predicted reduction in prevalence of early childhood undernutrition conditional on cumulative economic growth of 5, 16, 50 or 100 percent. The impact is measured in percentage points of the population under 3 years of age, for a country starting with a prevalence rate of 50 percent. The bottom row indicates the annual growth rates that would be required to cumulatively increase real GDP per capita by 5, 16, 50 or 100 percent between 2015 and 2030.