Identifying Catch-Up Trajectories in Child Growth: New Methods with Evidence from Young Lives

Other

Policy Research Working Paper 8353

Identifying Catch-Up Trajectories in Child Growth

New Methods with Evidence from Young Lives

Sam Jones
Jere R. Behrman
Hai-Anh H. Dang
Paul Anand

WORLD BANK GROUP
Development Economics
Development Data Group
February 2018

Policy Research Working Paper 8353

Abstract

Definitions of catch-up growth in anthropometric outcomes among young children vary across studies. This paper distinguishes between catch-up in the mean of a group toward that of a healthy reference population versus catch-up within the group, associated with a narrowing of the outcome distribution. In contrast to conventional empirical approaches based on dynamic panel models, the paper shows how catch-up can be tested via a latent growth framework. Combined with a flexible estimator incorporating individual-specific intercepts and slopes, this enables between- and within-group forms of catch-up to be tested in a unified setting. The application of the proposed approach reveals significant differences in the nature, extent, and drivers of catch-up growth across the four Young Lives countries (Ethiopia, India, Peru, and Vietnam). In addition, the paper shows how conclusions about catch-up are sensitive to the way in which anthropometric outcomes are expressed.

This paper is a product of the Development Data Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at sam.jones@econ.ku.dk or hdang@worldbank.org.

Identifying Catch-Up Trajectories in Child Growth: New Methods with Evidence from Young Lives

Sam Jones* Jere R. Behrman Hai-Anh H. Dang Paul Anand

*Jones (sam.jones@econ.ku.dk, corresponding author) is an Associate Professor in the Department of Economics, University of Copenhagen; Behrman is the William R. Kenan, Jr. Professor of Economics at the University of Pennsylvania; Dang (hdang@worldbank.org; corresponding author) is an economist in the Survey Unit, Development Data Group, World Bank, and a non-resident senior research fellow with Vietnam's Academy of Social Sciences; Paul Anand is a Professor at the Open University and Research Associate at HERC in Oxford University. We acknowledge helpful comments from Thomas Sohnesen, Hanna Szymborska, Jed Friedman and Ha Nguyen. Jones was funded by the Open University. Dang was funded by the Strategic Research Program (SRP) supported by UK Department of International Development.

1 Introduction

It is widely recognized that early childhood represents a critical period for the development of human capabilities Heckman (2006; 2007). Disadvantages experienced during this period can cast a long shadow, often affecting outcomes over the life course (Black et al. 2017). Among a range of risk factors, the effects of poor nutrition have received extensive attention, especially in developing country contexts (de Onis et al. 2012). While both the nature and average impacts of these risk factors are generally well known, differences in how children respond and cope with early disadvantage are not fully understood (Engle et al. 1996). A key area of debate concerns catch-up growth, which refers to recovery in anthropometric outcomes (height, weight) from initial disadvantage. Identifying family and community characteristics, as well as governmental policies, that systematically enhance resilience to shocks and promote developmental recovery remains an important research agenda (Almond et al. 2017, Anand et al. 2016).

Our point of departure in this study is that extant definitions and associated statistical tests for catch-up growth among children are not entirely convincing. This raises a fundamental challenge – advances in identifying relevant factors that promote recovery are likely to be limited if the presence (absence) of catch-up growth is not measured correctly. Indeed, until recently a prevailing view was that catch-up growth among stunted children was only possible during the child's first 24 months. In a landmark study, Adair (1999) tracked the heights of around 2,000 Filipino children from 2 to 12 years of age and found a significant reduction in the prevalence of stunting over time – i.e., recovery occurred beyond 2 years of age. However, on reviewing the same evidence, Cameron et al. (2005) conclude these results were driven by regression to the mean effects and, using alternative methods, find no evidence of catch-up growth in the same sample.

Methodological debates around catch-up growth remain alive (e.g., Hirvonen 2014). The aim of this paper is to revisit the ways in which catch-up growth is defined, as well as how it is measured in practice. We contribute to the literature in four ways. First, we propose a unified set of tests for catch-up growth. We carefully distinguish between two distinct forms of catch-up, namely: (i) convergence of initially disadvantaged individuals within a sampled population toward the group mean (within-group catch-up); and (ii) convergence of the group mean toward that of the median child from a healthy external reference population (between-group catch-up).[1] Existing studies in the literature have typically focussed on either one of these forms. However, since within-group catch-up has no necessary relation to between-group catch-up, an exclusive focus on any one form fails to provide comprehensive insights regarding recovery dynamics.

Our second contribution it to embed the two tests of catch-up in a latent growth framework. This

[1] The distinction between means and medians is important. Conventional methods to standardize anthropometric outcomes against a reference population employ the median as the relevant measure of central tendency. At the same time, least squares estimates of catch-up growth focus on means. So, if the mean for a standardized outcome (e.g., height-for-age z-score) in a sampled group approaches zero, then the average child in the sample corresponds to the median of the reference distribution. We retain the relevant distinction between means and medians throughout.

framework has been used to study child development before, but it has not been widely used as a basis to test for catch-up growth. We show how the framework allows both forms of catch-up to be tested in a unified setting. In addition, we apply a novel estimator that relies on minimal statistical assumptions regarding the data generating process. The merit of this procedure is demonstrated with longitudinal data from the four Young Lives countries – Ethiopia, India, Peru and Vietnam. On a technical level, we find the proposed fixed-effects estimator with individual slopes outperforms conventional correlated random effects (mixed effects) models. On the substantive level, the procedure reveals limited evidence for catch-up growth in either stature or body mass across the four countries.

Our third contribution is to verify the extent to which estimates of catch-up are sensitive to the ways in which anthropometric outcomes are transformed, often via standardization procedures. This point has been acknowledged before (e.g., Leroy et al. 2015), but few studies include a systematic comparison of the degree to which statistical tests of different forms of catch-up depend on the specific metric or transformation chosen. We find that externally-standardized z-scores and associated binary transformations (e.g., stunted / not stunted) tend to produce more positive assessments of the presence of catch-up growth compared to tests based on absolute differences. As this may be driven by changes in the variance of the reference distribution, we propose that the ratio between the outcome for a given child and the expected outcome for a child of the same age and gender in a reference population is used as a complementary measure.

Our final contribution is to extend the latent growth framework to show how the determinants of variation in mean child size and growth velocity can be investigated via a second stage of analysis. This provides a consistent and straightforward approach to identifying factors that account for differences in growth trajectories across children. As applied to the Young Lives data we find significant heterogeneity in the magnitude (and direction) of different drivers across the four countries. We also find that community characteristics explain a non-trivial share of the variance in child size and growth velocity.

The remainder of the paper is structured as follows: Section 2 reviews conventional definitions of catch-up and briefly summarizes recent literature. Addressing the shortcomings of extant empirical approaches, Section 3 sets out our proposed alternative procedure based on a latent growth framework; and we discuss alternative empirical estimation strategies. Section 4 applies our recommended approach to the Young Lives data and discusses the results. Section 5 concludes.

2 Identifying catch-up growth

2.1 Definitions

Existing literature on catch-up growth spans contributions across the medical, biological and social sciences.[2] Despite the huge volume of scholarship, studies do not employ a unique or consistent definition of what actually constitutes catch-up growth. Two main uses of the concept can be distinguished, each corresponding to different substantive research questions. The first focuses on whether initially disadvantaged children grow at a more rapid rate than other children in the sample population – i.e., they converge towards the mean of the group (Wit and Boersma 2002).[3] The second notion of catch-up concerns whether the average child in the sampled population is converging toward the median outcome observed in a healthy external reference population. So, in the first notion, interest is on how the *shape* of the outcome distribution evolves for a given sample over time. In the second notion the focus is on how the mean of this distribution evolves relative to the median child in a healthy population. Put differently, the first notion of catch-up focuses on within-group processes of convergence, while the second attends to between-group processes.

To appreciate the difference between the two notions of catch-up growth it is helpful to review what conventional empirical tests capture. Various studies of catch-up growth use some version of the following model:

$$y_{it} = \alpha + \beta y_{i,t-j} + x'_{is}\gamma + \varepsilon_{it} \qquad (1)$$
$$\text{where } s \in \{\emptyset, t, ..., t-j, ..., 0\}$$

where y is an anthropometric outcome of interest; $i \in \mathscr{I} = \{1, ..., I\}$ indexes individual children; t represents the child's age at the time of measurement; x is a vector of observed explanatory variables (e.g., household income); α is an intercept; ε is residual error; and subscript s is defined so as to encompass control variables observed at different points in time, as well as time-invariant characteristics.

Existing studies based on this model often focus on estimates for β. When outcomes are standardized by an external reference distribution, it is often argued that complete catch-up obtains when $\beta = 0$, and no catch-up obtains when $\beta = 1$. Setting all the elements of γ to zero and assuming only two periods $(t, t-j)$ for simplicity, it is easy to see that least squares estimates of equation (1)

[2] In recognition of this, throughout this paper we draw on studies from a wide range of fields. For now, we use the term 'growth' in reference to various child development outcomes of interest (e.g., height, weight, BMI etc.). We are specific later.

[3] Related studies in this vein also seek to identify specific factors that influence variation in individual growth rates *within* the given sample. See further below.

will yield the following slope coefficient estimate:

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_{i\in\mathscr{I}}(y_{i,t}-\bar{y}_t)(y_{i,t-j}-\bar{y}_{t-j})}{\sum_{i\in\mathscr{I}}(y_{i,t-j}-\bar{y}_{t-j})^2} \\
&= \frac{\sum_{i\in\mathscr{I}}(\tilde{y}_{i,t}\tilde{y}_{i,t-j})}{\sum_{i\in\mathscr{I}}(\tilde{y}_{i,t-j})^2} = \frac{1/I\sum_{i\in\mathscr{I}}(\tilde{y}_{i,t}\tilde{y}_{i,t-j})}{s^2_{t-j}} \\
&= \rho_{t,t-j}\frac{s_t}{s_{t-j}}
\end{aligned} \tag{2}$$

in which $\tilde{y}_{i,t} = (y_{i,t}-\bar{y}_t)$ is the demeaned outcome at time t; s_t is the sample standard deviation of y_t (calculated over all I sampled children observed at time t); and $\rho_{t,t-j}$ is the j^{th} order coefficient of autocorrelation.

In the case that the outcome (in each period) is standardized in relation to some reference distribution with median μ_t and standard deviation σ_t, we note that:

$$\begin{aligned}
& y^*_{i,t} = (y_{i,t}-\mu_t)/\sigma_t \\
\Longrightarrow\ & \tilde{y}^*_{i,t} = \frac{(y_{i,t}-\mu_t)-(\bar{y}_t-\mu_t)}{\sigma_t} = \frac{\tilde{y}_{i,t}}{\sigma_t}
\end{aligned} \tag{3}$$

In turn, the corresponding slope coefficient becomes:

$$\begin{aligned}
\hat{\beta}^* &= \frac{\sum_{i\in\mathscr{I}}(\tilde{y}_{i,t}\tilde{y}_{i,t-j})/(\sigma_t\sigma_{t-j})}{\sum_{i\in\mathscr{I}}(\tilde{y}_{i,t-j})^2/\sigma^2_{t-j}} \\
&= \frac{\sigma_{t-j}}{\sigma_t}\hat{\beta} = \frac{s_t}{\sigma_t}\frac{\sigma_{t-j}}{s_{t-j}}\rho_{t,t-j}
\end{aligned} \tag{4}$$

These expressions prompt three insights. First, both slope coefficient estimates are directly proportional to the autocorrelation coefficient, which captures the speed at which y_t converges towards its mean. Thus, the consistent component of this notion of catch-up is the (average) degree of persistence in the *divergence* of outcomes from the sample mean. Second, as equation (3) indicates, the slope coefficient estimates are unaffected by adding or subtracting a fixed constant in any period, such as μ_t – i.e., the β and β^* estimates are only related by the ratio of the scaling factors (standard deviations) applied in the standardization procedure. Consequently, the slope estimates in this formulation contain no direct information about the direction in which outcomes for the average child in the sample are moving. It follows that neither estimate for β can be informative about between-group catch-up, and this holds regardless of how the outcome is expressed, including whether it is standardized by an external reference distribution such as that provided by the WHO.

Third, despite the previous points, estimates of the slope coefficient will vary according to the standardization imposed. In the case where outcomes are standardized internally (i.e., $\sigma_{y_t} \equiv s_t$), then the slope coefficient *is* the autocorrelation coefficient. If raw outcomes are used, such as height, then the slope coefficient captures the autocorrelation coefficient multiplied by the ratio of the standard deviation of outcomes at time t to that in the previous period (see equation 2). For externally standardized outcomes, the scaling factor will be greater than one if the relative dispersion of the sample distribution to that of the reference population is growing over time. Consequently, in anything other than the internally standardized case, slope coefficients derived from equation (1) will not exclusively capture the speed of convergence toward the sample mean. In itself, this ambiguity motivates a search for alternative tests of this form of catch-up.

A second approach to identifying catch-up growth is often found in studies of disadvantaged populations (samples). In contrast to focussing on within-group convergence, Hirvonen (2014) recommends a shift of attention toward α in equation (1). Least squares estimates for this coefficient imply:

$$\hat{\alpha} = \bar{y}_t - \hat{\beta}\bar{y}_{t-j} - \vec{x}'_{is}\hat{\gamma} \tag{5a}$$

$$\hat{\alpha}^* = (\bar{y}_t - \mu_t) - \hat{\beta}(\bar{y}_{t-j} - \mu_{t-j}) - \vec{x}'_{is}\hat{\gamma} \tag{5b}$$

and where the latter estimate refers to externally-standardized outcomes. In this case, the same author suggests the simple difference captures between-group catch-up, defined as the extent of convergence of the sample mean toward that of the reference population. In the unconditional case, this is obtained directly from equation (5b) when $\gamma = 0 \wedge \beta = 1$, which is equivalent to examining mean differences in the standardized outcome, $\Delta\bar{y}^*$. Similarly, Cameron et al. (2005) argue that $\hat{\alpha}$ provides a meaningful measure of catch-up in stature from initial stunting. However, they recommend β remains unconstrained in order to address potential regression to the mean effects (see further below; also Adair 1999).

2.2 Existing literature

The above discussion distinguished between two forms of catch-up growth. Table 1 summarizes a selection of ten recent papers in both these traditions (published in the past 5 years), all of which refer to developing country contexts.[4] From this, four general comments merit note. First, existing studies have predominantly considered catch-up growth in relation to child height. One rationale is that low height-for-age is associated with stunting, which is known to adversely affect outcomes in adulthood (Almond et al. 2017, Hoddinott et al. 2013). However, low BMI-for-age values are associated with child wasting, which is a symptom of acute under-nutrition. Shrimpton et al. (2001) argue that adverse *in utero* conditions tend to affect child birth weight more than length, but in contrast to height disadvantages, weight disadvantages can be addressed relatively more quickly in

[4] The studies were identified from Google Scholar based on the combined search text: ‘ "catch-up growth" & children & "developing countries" ’. The search period was limited to 2012-2017; only primary empirical studies undertaken in developing countries were retained; and those concerned with catch-up growth either after specific interventions or associated with specific medical conditions were excluded. Papers were also excluded that have a primary focus on the determinants of variation in growth rates (e.g., Georgiadis et al. 2016; 2017). The final list (in the table) includes the most relevant peer-reviewed studies; but we recognize this is somewhat subjective and is likely to be selective.

Table 1: Summary of recent studies of catch-up growth

Study	Context	Outcomes	Metric(s)	Findings
M2012	Indonesia	Height	β	Incomplete within catch-up
O2013	Ethiopia	HAZ	β	Incomplete within catch-up
S2013	Ethiopia, India, Peru, Vietnam	HAD	β	Some within-catch-up, especially before age 5
F2014	Ethiopia, India, Peru, Vietnam	HAZ	$\Delta\bar{y}$, β	Some within-catch-up, no between catch-up
L2014	Brazil, Guatemala, India, Philippines, RSA	HAZ, HAD	$\Delta\bar{y}$	Some between catch-up (HAZ), persistent deficits (HAD)
L2015	Ethiopia, India, Peru, Vietnam	HAZ, HAD	$\Delta\bar{y}$	No between catch-up (HAZ & HAD)
T2015	Malawi	HAZ, HAD	$\Delta\bar{y}$	Some between catch-up (HAZ), persistent deficits (HAD)
H2016	China, South Africa, Nicaragua	HAZ	β	Incomplete within catch-up
Z2016	Bolivia (Amazon)	Height, HAZ	β	Incomplete within catch-up, persistent deficits
S2017	Timor-Leste	HAZ, BAZ	$\Delta\bar{y}$	Some between catch-up in height (HAZ), not BMI (BAZ)

Notes: Studies are classified according to the main parameter on which analysis focuses, which we (re)interpret according to the distinction between within- and between-group convergence; 'Metrics' are the parameters and refer to equation (1), where $\Delta\bar{y}$ is the unconditional change in the mean sample outcome; HAZ and HAD are height-for-age z-scores and absolute deficits respectively, measured relative to a well-nourished external population; BAZ is the BMI-for-age z-score, also measured relative to a well-nourished external population; RSA denotes South Africa.

Study abbreviations: M2012 = Mani (2012), O2013 = Outes and Porter (2013), S2013 = Schott et al. (2013), F2014 = Fink and Rockers (2014), L2014 = Lundeen et al. (2014b), L2015 = Leroy et al. (2015), T2015 = Teivaanmäki et al. (2015), H2016 = Handa and Peterman (2016), Z2016 = Zhang et al. (2016), S2017 = Spencer et al. (2017).

the first few years of life, where conditions are appropriate. As such, consideration of trends in both height and BMI may provide valuable complementary insights. This is confirmed by Spencer et al. (2017), who find a divergence in trends for HAZ and BAZ scores (BMI-for-age z-scores) in two rural communities in Timor-Leste, calculated with reference to WHO (World Health Organization) charts.[5]

Second, the vast majority of studies focus on only one or the other form of catch-up growth. While this is reasonable in specific cases, it means evidence on catch-up growth (broadly defined) is often partial, and little is known about the relationship *between* the two forms of catch-up. This is important because the forms of catch-up are distinct and bear no necessary (logical) relation to

[5] The main research goal of Spencer et al. (2017) is to calculate community-specific growth curves. However, their comparison of these curves against WHO standards provides an indication of between-group catch-up.

one another. Admittedly, a number of the studies include an analysis of recovery from stunting, which is identified via a binary transformation of HAZ scores.[6] For instance, Teivaanmäki et al. (2015) compare the evolution of HAZ scores for children classed as not stunted, moderately stunted and severely stunted at baseline. They find that HAZ scores improved for children in each class, which is indicative of between-group convergence. But they also find that HAZ scores improved at a somewhat faster pace for the severely stunted, which would be consistent with the presence of within-group catch-up (see also Fink and Rockers 2014, Zhang et al. 2016). While these kinds of stratified analyses do touch on the different forms of catch-up, their findings are rarely systematized so as to shed light on the correspondence between alternative forms of convergence.

Third, our review of existing studies reveals there is no unified framework for testing both forms of catch-up – i.e., the two forms are analyzed separately, often via distinct models. This raises doubts as to whether the findings are mutually consistent. For instance, estimates of between-group catch-up tend to focus on average changes in outcomes. In the case of conditional growth models, these use estimates for α in equation (1) derived under the constraint that $\beta = 1$, which rules out estimation of the autocorrelation coefficient.[7] More generally, the use of multiple empirical operationalizations of catch-up growth makes it difficult to consistently interpret and compare results across different studies.

Fourth, the papers referenced in Table 1 predominantly rely on estimates for catch-up based on some form of equation (1) (as do many earlier studies; e.g., Adair 1999, Hoddinott and Kinsey 2001). However, these estimates are prone to bias. A primary concern stems from measurement error in the outcome variable. Assuming the error in each measurement period is white noise, such that the observed outcome includes the true component plus error: $y_{it} = y^*_{it} + \nu_{it}$, it can be shown that least squares estimates of β will be biased towards zero to an extent proportional to the magnitude of measurement error in period $t - j$. This is a main form in which regression to the mean effects arise (Lohman and Korb 2006, Tu and Gilthorpe 2007; 2011) and is pertinent in the current context since anthropometric measurement error may be commonplace and relatively more acute among younger children (e.g., where exact age is not known; see Ulijaszek and Kerr 1999, Zhang et al. 2016).[8] Such effects are also a major concern where units are stratified by baseline outcomes (Cameron et al. 2005, Jerrim and Vignoles 2013) – e.g., the true rate of recovery from initial stunting may be inflated by regression to the mean effects. Furthermore, by the mechanics of least squares, any bias in β will directly effect estimates for α.

An additional source of bias is due to the omission of variables that are correlated with the lagged outcome or the included set of controls (x_{it}). The latter play two important roles. First they can be used to control for idiosyncratic deviations in the composition of the sample, either over time

[6] Typically, a child with a height-for-age z-score below -2 is classified as stunted.
[7] This is distinct from studies that estimate equation (1) where the lagged outcome is subtracted from both sides, the reason being that in these cases it is the estimates for β that remain in focus.
[8] A related challenge is noise induced by differences in the timing of growth spurts. These may be captured by richer models – see further below.

or relative to the complete sample frame, in turn reducing any bias in the parameter estimates deriving from compositional sources (e.g., an excess number of wealthier children in the actual sample). Second, they can be used to study heterogeneity in growth patterns – e.g., via inclusion of interaction terms or through auxiliary residual analysis (see Section 4.4). If relevant elements of x_{it} have been omitted then conditional estimates of catch-up may be biased, or our understanding of growth heterogeneity will be limited.

In the context of longitudinal data, a promising response is to include individual fixed effects (Handa and Peterman 2016). Typically these account for a large amount of the variation in the data and are able to control for all time-invariant factors (at the individual level) that influence growth outcomes. However, the dynamic nature of the specification makes inclusion of such terms highly problematic. Panel estimates of equation (1) correspond to a dynamic or autoregressive panel situation in which the hypothesis of no catch-up (convergence) implies the outcome follows a random walk. A large econometrics literature deals with the technical problems of rigorously identifying the parameters of dynamic panel models in the presence of (unobserved) heterogeneity. Bond et al. (2005) review the properties of equation (1) under the simplifying assumption that the error term follows: $\varepsilon_{it} = (1-\beta)\alpha_i + v_{it}$, where α_i is an individual-specific constant and v_{it} is white noise. They show that standard OLS estimates of (1) are consistent under the null hypothesis of a unit root ($\beta = 1$); but estimates for β will be upward biased under the alternative (see also Baltagi 2008, Bond 2002). They also show that popular solutions to this problem, such as those based on instrumental variables or GMM techniques, are not a panacea. While alternative estimators can be consistent, they often have low power in finite samples. Furthermore, and as is well known, finding instruments that are valid and perform well across different contexts is a formidable task.

The challenges associated with addressing unobserved individual heterogeneity in the dynamic panel context have not been widely appreciated in empirical work on catch-up growth. Rather, individual (unobserved) heterogeneity is often either ignored, or is swept out by first differencing (Handa and Peterman 2016, Mani 2012). This contrasts with research in other fields where these issues are often placed at center stage – e.g., as in empirical tests of the relationship between initial firm size and subsequent growth (e.g., Ribeiro 2007), as well as in the literature on aggregate growth (e.g., Durlauf et al. 2005). Furthermore, individual heterogeneity in child growth patterns is not only highly plausible, such as due to significant differences in initial conditions (e.g., birth weight and length), but also the ways in which time-invariant child characteristics affect subsequent growth are often of stand-alone interest. Thus, tests for catch-up that are able to take into account individual heterogeneity – not only as an additive control but also as a determinant of variation in growth patterns – are called for.

3 Latent growth methods

The previous section argued that existing studies of catch-up growth are limited in scope and often unconvincing. Critically, a number of the main sources of potential bias affecting estimates of equation (1) are difficult to address within a dynamic panel framework. Latent growth models provide an alternative means to analyze processes of child development. Models of this sort are widely applied to understand child development, often to quantify and explain (socio-economic) gradients in growth outcomes.[9] In this section we show how latent growth models also provide a unifying framework for identifying the nature and extent of catch-up growth, addressing the shortcomings associated with the conventional framework (equation 1). In addition, we recommend a flexible empirical estimation strategy that relies on weaker assumptions than typically imposed by conventional random or mixed effects estimators.

A general latent growth framework, which applies naturally to child growth, is as follows:

$$y_i(t_i) = (\alpha_0 + \alpha_i) + (\beta_0 + \beta_i) f(t_i) + x_{it}'\gamma + \varepsilon_{it} \tag{6a}$$

$$\text{Var}(\alpha_i) = \sigma_\alpha^2,\ \text{Var}(\beta_i) = \sigma_\beta^2,\ \text{E}(\alpha_i\beta_i \mid x_{it}) = \rho_{\alpha\beta}(\sigma_\alpha\sigma_\beta) \tag{6b}$$

$$\text{E}(x_{it}'\alpha_i \mid \beta_i) = \rho_{x\alpha}(\sigma_\alpha\sigma_x),\ \text{E}(x_{it}'\beta_i \mid \alpha_i) = \rho_{x\beta}(\sigma_x\sigma_\beta) \tag{6c}$$

Notation here is not exactly the same as before. Consistent with our interest in both between- and within-group catch-up, y must be referenced to an external population. Also, the nature and interpretation of α and β are altered. Both are comprised of two parts: first, a sample average component, taking a zero subscript; and, second, an unobserved individual-specific component, taking the i subscript, which are mean zero by construction.[10] The α terms capture the level of the outcome variable when $f(t_i) = 0$, where t represents the child's age and $f(\cdot)$ is an unspecified functional form. The β terms represent the velocity of growth – i.e., the expected change in the outcome for each unit increment in age. Thus, in the terminology of Tanner (1962), the two unobserved components correspond to estimates of deviations from mean child size and growth velocity respectively. Lastly, to facilitate interpretation of the α terms, the set of time-varying controls is demeaned.

A latent growth framework is consistent with theoretical models of early child growth (see De Cao 2015). It also holds a number of advantages in comparison to dynamic panel specifications. First, the framework permits simultaneous examination of a set of distinct null hypotheses about different forms of catch-up growth. Estimates for the sample average slope β_0 capture the mean growth velocity in the sample. Thus, where the outcome is expressed relative to a healthy reference population (e.g., as per externally-referenced HAZ scores), positive estimates indicate a faster pace

[9] A recent example in this vein is McCrory et al. (2017). Also see Georgiadis et al. (2016) for similar analysis, using a path model.

[10] Estimation of individual heterogeneity in the growth trajectories is the 'latent' term in the description of the approach.

Table 2: Alternative combinations of hypotheses regarding catch-up growth

		Between convergence? Yes	Between convergence? No
Within convergence?	Yes	$(\alpha_0 < 0 \wedge \beta_0 > 0) \wedge \rho_{\alpha\beta} < 0$	$(\alpha_0 \geq 0 \vee \beta_0 \leq 0) \wedge \rho_{\alpha\beta} < 0$
	No	$(\alpha_0 < 0 \wedge \beta_0 > 0) \wedge \rho_{\alpha\beta} \geq 0$	$(\alpha_0 \geq 0 \vee \beta_0 \leq 0) \wedge \rho_{\alpha\beta} \geq 0$

Note: '$\wedge$' is the logical AND operator; '$\vee$' is the logical OR operator; these conditions assume outcomes are expressed relative to a healthy external reference population and a positive change in the outcome always corresponds to an improved situation.

of growth relative to the reference group. Consequently, if average (or initial) size is below that of the reference group, then a positive mean slope would indicate population-wide catch-up has occurred. Within-group catch-up (convergence) is captured by the relationship between child size and velocity ($\rho_{\alpha\beta}$). If the correlation is negative, then below-average children are growing faster than their counterparts and the predicted distribution of outcomes is converging toward a common mean over time. If the correlation is positive, above-average children are extending their advantage and the predicted outcome distribution is widening over time. Thus, within-population or relative catch-up requires: $\rho_{\alpha\beta} < 0$, and where the latter correlation is conditional on the included controls (x_{it}).

Second, unlike the conventional framework, there is no lagged outcome in the set of explanatory variables. As a result, measurement error in the outcome variable will only reduce the overall explanatory power of the model, but it should not bias estimates of the catch-up parameters. Third, the absence of the lagged outcome variable means that the unobserved individual intercepts and slopes can be retrieved directly (see below), side-stepping the various problems encountered in dynamic panels. Fourth, the framework easily extends to multiple observation periods and, where data permit, allows for non-linearities in growth patterns to be estimated and tested via choices for $f(t_i)$ (e.g., Chirwa et al. 2014).[11]

Based on this framework, Table 2 sets out specific hypotheses regarding the combinations of convergence that may be identified. Formally, absolute convergence of the sample mean toward the median of a healthy external reference population requires: $\alpha_0 < 0 \wedge \beta_0 > 0$. The appropriate null hypotheses to be taken to the data, corresponding to an absence of either form of catch-up, are located in the bottom right cell of the table '(No, No)'. In the present context, the analytically interesting alternative to the null hypotheses is that catch-up growth *is* occurring. As a result, we state the null hypotheses using inequality signs, implying that one-sided statistical tests are appropriate. Put differently, conventional tests of a general null hypothesis that, say, $\alpha_0 \neq 0$ are not particularly informative about catch-up growth.

The outstanding issue is how these hypotheses might be tested. This requires selection of an

[11] In our application (Section 4) we have four observations per individual so non-linear functional forms cannot be considered. Rather, we use the simple linear function: $f(t_i) = t_i - \bar{t}$.

econometric approach, yielding parameter estimates and standard errors, as well as a method to combine tests of individual parameters into composite hypotheses about catch-up. With respect to the first question, existing approaches to estimation of growth trajectories correspond to special cases of equation (6a), distinguished by specific restrictions imposed on the moments given by equations (6b) to (6c). The most simple cross-section approach restricts all moments to zero, which amounts to setting $\forall i : \alpha_i = 0 \wedge \beta_i = 0$, and treats all forms of unobserved individual heterogeneity as a nuisance. While this can be estimated easily via pooled OLS (POLS), many scholars consider these restrictions excessively strong. Moreover, for present purposes, a POLS estimator does not permit direct tests of the full set of catch-up hypotheses (Table 2).

An established alternative to pooled OLS is a mixed effects estimator, in which the unobserved heterogeneity terms (α_i, β_i) are treated as random effects, presumed to follow a multivariate normal distribution. Specific instances of these models also permit the unknown correlation $\rho_{\alpha\beta}$ to be estimated. Such Correlated Random Effects (CRE) estimators are a popular choice and allow the full range of catch-up hypotheses to be tested (see Johnson 2015). Despite these attractions, in most applications of mixed effects models the final two moments of the general system (equation 6c) remain restricted to zero – i.e., all elements of x are treated as orthogonal to the unobserved heterogeneity terms. In part, this assumption derives from technical challenges of convergence – the more parameters there are to be estimated, model convergence takes significantly longer and is more prone to failure (Chirwa et al. 2014, Gurka et al. 2011). The trade-off is that misspecification of the structure of correlation between so-called fixed and random effects is not innocuous and will deliver inconsistent estimates where independence restrictions are violated (see Hausman 1978, Jacqmin-Gadda et al. 2007).

In light of these considerations, a further option is to treat the unobserved heterogeneity terms as 'fixed' latent variables. While fixed-effects estimators for additive terms are commonplace and generally preferred within the economics literature, estimators allowing for fixed individual-specific slopes are less known. The latter models, often referred to as FEIS estimators (fixed effects with individual slopes), were originally introduced by Polachek and Kim (1994) but were not widely disseminated mainly due to implementation challenges. In our case, the advantage is that FEIS estimators impose no restrictions on any of the moments, nor do they impose specific distributional assumptions on the latent variables. Furthermore, recent advances in computing power and estimation algorithms mean these models can now be implemented with relative ease, even in the context of large panels (Guimarães and Portugal 2010). We therefore review the performance of this class of estimators, which to our knowledge have not been previously applied to analyze latent growth.

Appendix Table A1 summarizes some of the primary advantages and disadvantages associated with alternative econometric approaches to estimating the latent growth model. It notes that although the FEIS model is a flexible alternative to CRE models, it does introduce additional costs. First, we expect a reduction in efficiency (loss of degrees of freedom) since a large number of additional parameters must be estimated. As such, the gain in fit of the FEIS model, relative to alternatives, merits examination. Second, under the FEIS estimator, all time-invariant elements of x are absorbed by the latent variables. This naturally focuses attention on the predicted effects of age-specific changes in the elements of x (within-unit variation), holding each individual's mean size (intercept) and growth rate (slope) fixed. As Bell and Jones (2015) explain, there is no necessary reason why the effect of within-unit variation in x on the outcome will always be the same as the effect of between-unit variation, one reason being that the latter reflects conditions at a higher level of aggregation. In longitudinal individual data, the distinction of within- versus between-unit variation effectively captures the difference between transitory (short-run) and permanent (long-run) effects. So, applying an FEIS estimator to equation (6a) would effectively address bias from unobserved static heterogeneity. But the downside is these estimates do not yield direct insights into the effects of long-run elements of x on the outcome, including time-invariant characteristics. This drawback is not crucial for testing the basic hypotheses about catch-up. Even so, in Section 4.4 we show how further insight can be gained by examining the correlates of the estimated fixed effects in a second stage analysis.

The remaining issue is that hypotheses about catch-up (Table 2) are composite in nature. In order to evaluate claims about catch-up, we must combine results from various individual hypothesis tests, while also minimizing type I and II errors. Following Wilkinson (1951), the maximum probability (maxP) associated with each individual null hypothesis represents a straightforward omnibus test of whether they can be rejected in all cases simultaneously (i.e., it is a test of their disjunction). We use this approach here.[12] To address type II errors from multiple testing, we further adjust the confidence level using the conservative Bonferonni procedure, which is robust to arbitrary dependence between the individual tests (Clarke and Hall 2009). So, to test for between-group convergence, we take the highest probability of the separate one-sided nulls ($\alpha_0 \geq 0; \beta_0 \leq 0$) and adjust the chosen confidence level by a factor of two. Similarly, to test for both within- and between-convergence, we take the highest probability of the three one-sided nulls and adjust the chosen confidence level by a factor of three.

4 Application to Young Lives

The rest of this paper applies the proposed approach to the rich longitudinal data on child development collected under the Young Lives (YL) initiative, which covers four low-income countries.[13] Section 4.1 briefly introduces the data; Section 4.2 reports the baseline results for tests of the

[12] Various alternative omnibus tests have been proposed. However, many of these are designed for cases where the number of hypotheses to be tested (jointly) is very large (e.g., Roback and Askins 2005). Also, various other omnibus tests, such as Fisher's combined probability test, essentially provide a kind of average probability and therefore can lead to a rejection of the null even if some individual probabilities (hypotheses) are insignificant.

[13] The YL data have been used extensively before, therefore detailed introduction is not necessary. For further details and examples see (*inter alia*) Barnett et al. (2012), Crookston et al. (2013), Lundeen et al. (2014a).

catch-up hypotheses (Table 2), focussing on conventional HAZ scores; Section 4.3 considers the sensitivity of these results across alternative metrics of height as well as BMI; and Section 4.4 explores specific time-invariant child characteristics associated with child size and growth velocity.

4.1 Data

We use the four latest public-use rounds of the YL data and focus exclusively on the youngest cohort, who were approximately one year old in the first round and around 12 years old in the fourth round.[14] Table 3 summarizes the analytical dataset we use hereafter. We have removed children who have missing data for their age (1,116 observations removed), or who were under 6 months old in the first round (165 observations removed). Also, due to the longitudinal nature of the intended analysis, we excluded children observed in fewer than three of the four rounds (452 observations removed). This yields a final dataset containing 30,515 observations on 7,681 children.[15] As Part (a) of the table shows, each round contains data on almost 2,000 children in all four countries. The table also indicates that the sample is well balanced by gender, and there are only small differences in the ages of sampled children (given as the age in months divided by 12) across countries in each collection round.

The descriptive statistics in Table 3(a) report averages across rounds for a range of growth outcomes. Child height (in cm) and weight are measured directly, from which the BMI (body mass index) is derived.[16] Notably, average heights are broadly comparable across the countries, but mean BMIs differ – e.g., contrast Peru (17.6) and India (14.7). All remaining outcomes are expressed in relation to expectations for a child of the same age and gender in a healthy external reference population. For this, we use the latest WHO child growth standards and follow Vidmar et al. (2004; 2013) to estimate relevant healthy population medians and standard deviations.

Denoting the observed outcome (i.e., height, BMI) for child i at age t as y_{it}, and the corresponding median expected outcome estimated from the reference distribution as $\bar{\theta}_t$, we construct the following indicators: (i) standardized differences, $(y_{it} - \bar{\theta}_t)/\sqrt{\text{Var}(\theta_t)}$, which yield conventional z-scores

[14] To enhance replication, we use the merged or constructed YL datasets, available at: `https://www.younglives.org.uk/content/young-lives-rounds-1-4-constructed-files`. The older cohorts are excluded since academic analysis of growth recovery (faltering) has typically focussed on the growth trajectories of younger children (e.g., Leroy et al. 2014, Victora et al. 2010). Furthermore, children from the older cohorts attain puberty during the data collection period, which is associated with a rapid spurt in growth. As such, the simplifying assumption of a linear (individual) growth trajectory, which is necessary given we have a maximum of four observations per child, is problematic for older cohorts. For the younger cohort, a small share (less than one-third) of girls have reached menarche by the fourth round (for further discussion see Schott et al. 2017).

[15] Given the structure of the model, children observed in just three rounds can be retained in straightforward fashion.

[16] We follow Crookston et al. (2013), among others, and adjust the raw heights and BMIs collected in the first round to account for differences in ages at the time of observation. As Victora et al. (2010) note, HAZ patterns in disadvantaged communities tend to decline sharply from birth until around 20 months, after which a more gradual pattern of change is found. This correction therefore addresses any non-linearities in the relation between outcomes and child age in the first round, and strengthens the general assumption of a linear trend in outcomes as applied in our regression estimates. However, this does not meaningfully alter our results.

Table 3: Descriptive statistics, by country

	Ethiopia		India		Peru		Vietnam	
	Mean	(st.err.)	Mean	(st.err.)	Mean	(st.err.)	Mean	(st.err.)
(a) Averages (Rounds 1 to 4):								
No. children	1,882	(0.08)	1,920	(0.11)	1,913	(0.30)	1,914	(0.23)
Child's age (years)	6.6	(0.05)	6.6	(0.05)	6.5	(0.05)	6.6	(0.05)
Female (%)	47.2	(0.58)	46.4	(0.57)	49.7	(0.57)	48.7	(0.57)
Household size	6.0	(0.02)	5.3	(0.02)	5.5	(0.02)	4.7	(0.02)
Height	109.6	(0.30)	108.8	(0.29)	109.4	(0.30)	110.5	(0.31)
Height z-score (HAZ)	-1.4	(0.01)	-1.4	(0.01)	-1.3	(0.01)	-1.1	(0.01)
Height ratio (HAR, %)	-5.8	(0.06)	-6.1	(0.05)	-5.3	(0.05)	-4.7	(0.05)
Height difference (HAD)	-6.9	(0.07)	-7.3	(0.07)	-6.1	(0.07)	-5.6	(0.07)
Not stunted (%)	71.0	(0.52)	69.9	(0.52)	74.7	(0.50)	79.2	(0.46)
Not severely stunted (%)	90.8	(0.33)	93.3	(0.29)	93.6	(0.28)	96.0	(0.22)
BMI	14.9	(0.02)	14.7	(0.03)	17.6	(0.03)	15.8	(0.03)
BMI z-score (BAZ)	-1.1	(0.01)	-1.2	(0.01)	0.6	(0.01)	-0.5	(0.01)
BMI ratio (BAR, %)	-8.9	(0.12)	-9.7	(0.17)	8.1	(0.16)	-3.4	(0.14)
BMI difference (BAD)	-1.5	(0.02)	-1.6	(0.03)	1.3	(0.03)	-0.6	(0.02)
BMI not low (%)	69.3	(0.53)	64.9	(0.54)	98.3	(0.15)	85.6	(0.40)
Healthy BMI (%)	62.1	(0.56)	60.9	(0.56)	66.2	(0.54)	75.8	(0.49)
(b) Long differences (Round 4 - Round 1):								
No. children	-19	(0.00)	-23	(0.00)	-70	(0.00)	-46	(0.00)
Child's age (years)	11.1	(0.01)	11.0	(0.01)	10.9	(0.01)	11.2	(0.01)
Female (%)	0.0	(0.00)	0.0	(0.00)	0.0	(0.00)	0.0	(0.00)
Household size	0.1	(0.06)	-0.6	(0.06)	-0.5	(0.06)	-0.4	(0.04)
Height	69.5	(0.16)	68.3	(0.16)	71.2	(0.15)	71.9	(0.17)
Height z-score (HAZ)	0.0	(0.04)	-0.2	(0.03)	0.4	(0.02)	0.0	(0.02)
Height ratio (HAR, %)	-1.9	(0.13)	-2.4	(0.11)	-0.1	(0.10)	-1.3	(0.10)
Height difference (HAD)	-6.4	(0.15)	-6.7	(0.15)	-3.6	(0.15)	-4.5	(0.16)
Not stunted (%)	7.8	(1.29)	-0.5	(1.23)	10.9	(1.06)	0.0	(1.00)
Not severely stunted (%)	11.6	(0.97)	3.7	(0.75)	6.0	(0.70)	-0.5	(0.58)
BMI	-1.0	(0.05)	0.6	(0.11)	1.9	(0.07)	1.1	(0.06)
BMI z-score (BAZ)	-1.2	(0.04)	-0.3	(0.03)	-0.2	(0.03)	-0.2	(0.03)
BMI ratio (BAR, %)	-12.1	(0.30)	-2.7	(0.61)	3.7	(0.43)	-1.1	(0.35)
BMI difference (BAD)	-2.2	(0.05)	-0.6	(0.11)	0.7	(0.07)	-0.2	(0.06)
BMI not low (%)	-36.9	(1.40)	-13.9	(1.40)	0.8	(0.44)	-13.4	(1.04)
Healthy BMI (%)	-19.2	(1.58)	-15.8	(1.48)	8.5	(1.48)	-18.2	(1.31)

Note: aside from raw height and BMI (body mass index), all anthropometric outcomes are calculated with reference to a healthy external population; stunting is given by a HAZ score below -2; severe stunting is given by a HAZ score below -3; a low BMI is below the 5th percentile of the reference distribution; a healthy BMI lies between the 5th and 85th percentiles; see text for further description of height and BMI transformations.
Source: own estimates.

(e.g., the height-for-age z-score, HAZ; BMI-for-age is denoted BAZ); (ii) raw difference scores, $y_{it} - \bar{\theta}_t$, yielding a metric of absolute divergence in outcomes; and (iii) the ratio of outcomes, $y_{it}/\bar{\theta}_t - 1$. Although not employed widely, the latter outcome ratios address a concern that absolute differences can widen over time, even if growth rates are held constant. Formally, if *initial* values on an outcome for two individuals strictly differ $y_{i0} > y_{j0}$ but grow at the same rate ($g > 0$) thereafter, the absolute difference, $\Delta_t = (y_{i0} - y_{j0})(1+g)^t$, will increase with t. In contrast, the ratio y_{it}/y_{jt} will be constant, reflecting only the initial relative difference. Consequently, the behavior of the ratio over time is expected to be more stable and meaningful.

A reason for expressing the same underlying outcomes in different ways is to investigate the sensitivity of measures of catch-up to alternative formulations (for elaboration see Section 4.3). Part (b) of Table 3 reports sample averages of long differences between the final and first rounds (calculated in pairwise fashion over each child). As expected, all countries display significant gains in mean raw height, ranging from 68 to 72 cms over around 11 years. However, the remaining height metrics paint a diverse picture. Both the difference and ratio metrics (HAD and HAR) show a deterioration in outcomes over time relative to the reference population in all four countries. In contrast, the standardized height metric (HAZ) points to small improvements in all countries other than India. The metrics of stunting, here valued positively, are based on standard HAZ thresholds. They show either improvement or no significant change across all countries over the four rounds.

Turning to the BMI scores, we note that in contrast to the other three countries, Peru's mean standardized score (BAZ) is significantly greater than that of the reference population, implying children in the sample are heavier than expected (given their height and age). The long differences for the BMI scores also reveal the Peruvian sample to be an exception. While all other countries tend to show some deterioration in their BMI scores over time relative to the healthy external population (regardless of the indicator chosen), the same scores in Peru tend to increase. The only exception is the BAZ score for Peru, which records a small but not statistically significant decline.

For reference, Appendix Figures A1 to A6 illustrate trends in the various height and BMI metrics for the four countries over the four rounds. Each figure plots the predicted linear trend (by age) for the 5th percentile, the median and the 95th percentile of the score distributions. The figures confirm the overall pattern in aggregate outcomes described in Table 3(b). However, they additionally reveal important differences in how the shape of the distributions appear to be evolving. For example, the HAZ plots (Figure A1) for Ethiopia and India indicate the distribution is narrowing over time (with age), which would be consistent with a process of within-group convergence in z-scores. However, the corresponding HAD plots (Figure A3) reveal clear distributional divergence; and the HAR plots (Figure A2) indicate more moderate change – slight convergence in the case of Ethiopia and slight divergence in India. Charts for the BMI measures also indicate significant heterogeneity across the countries. We explore these differences further in Section 4.3.

4.2 Baseline results

Our baseline results concerning catch-up growth in the four YL countries are summarized in Table 4, which takes height-for-age z-scores (HAZ) as the primary outcome of interest.[17] Results for three different estimators of equation (6a) are reported. In each case, the same specification is employed and we include a set of control variables spanning a range of both fixed and time-varying characteristics. These are *inter alia*: the child's gender; whether children come from the majority ethno-linguist group (ELG) of the region of residence; whether they have older siblings; if the location is urban; the age and highest level of education of their mother (at time of birth); a household wealth index; household size; and household exposure to shocks.[18] For comparison, Appendix Table A2 reports the same results, excluding the additional control variables. These are highly similar and thus do not merit separate discussion.

Before reviewing the results in detail, the scaling of the age variable merits comment. It is well known that estimates of both the intercept and the slope-intercept covariance in latent growth models are sensitive to the point at which one sets $t = 0$ (Biesanz et al. 2004, Stoel and Van Den Wittenboer 2003). While different choices can help address different analytical questions, Tu and Gilthorpe (2011) demonstrate that age/time must be demeaned (i.e., one must use $t_i^* = t_i - \bar{t}$) to avoid distortion of the null hypothesis of zero correlation between the slope and the intercept (see also Blance et al. 2005, Tu and Gilthorpe 2007). We apply this adjustment here, thus ensuring a clean interpretation of the estimated slope-intercept correlation coefficient under the null hypothesis that it is equal to zero. Consequently, the individual intercept terms $(\alpha_0 + \alpha_i)$ represent the expected outcome for the child at around 6.5 years of age.

For each country and estimator, Table 4 reports estimates for the three parameters of interest: the mean outcome in the sample (i.e., the average difference in stature relative to the reference median, α_0), the mean growth velocity (β_0), and the conditional correlation between the estimated latent variables ($\rho_{\alpha\beta}$). In the case of the POLS estimator, the latent variables are ignored and thus not estimated. Even so, a proxy for individual-specific slopes can be derived by interacting the child's age with the set of time-invariant controls. In turn, the correlation between these slopes and the estimated conditional mean size of each child yields a crude first-pass approximation to $\rho_{\alpha\beta}$. For the CRE estimator, the same correlation coefficient is estimated directly as a model parameter. For the FEIS estimator, implemented in Stata via the `reghdfe` command, the fixed effects are retrieved after estimation (see Guimarães and Portugal 2010). In turn, the correlation coefficient is calculated directly from the estimated fixed effects.[19]

[17] HAZ scores are chosen as they are the *de facto* standard in the literature (e.g., see Table 1).

[18] All control variables are demeaned (at the country-level) to assist comparison of estimates. The variables are chosen following previous studies and in accordance with variables that are common to all countries and rounds of the YL datasets. For further discussion see Section 4.4.

[19] To adjust for measurement error, the fixed effects are shrunk toward the sample mean of zero in accordance with the number of observations used to estimate each effect. This corresponds to an empirical Bayes procedure (see the discussion in Koedel et al. 2015), yielding a shrinkage factor of 80% in the majority of cases where the child is

Table 4: Estimates of catch-up growth in HAZ scores (conditional model)

					Between			Within		Both
		α_0	β_0	$\rho_{\alpha\beta}$	$H0_{\alpha_0}$	$H0_{\beta_0}$	?	$H0_{\rho_{\alpha\beta}}$	?	?
Ethiopia	POLS	-1.390	-0.016	-0.246	0.00	0.99	N	0.00	Y	N
		(0.022)	(0.004)	(0.023)						
	CRE	-1.402	-0.002	-0.529	0.00	0.69	N	0.00	Y	N
		(0.022)	(0.004)	(0.023)						
	FEIS	-1.381	0.003	-0.303	0.00	0.20	N	0.00	Y	N
		(0.023)	(0.003)	(0.023)						
India	POLS	-1.438	-0.034	0.367	0.00	0.99	N	0.99	N	N
		(0.020)	(0.004)	(0.023)						
	CRE	-1.451	-0.025	-0.295	0.00	0.99	N	0.00	Y	N
		(0.020)	(0.003)	(0.023)						
	FEIS	-1.440	-0.015	-0.193	0.00	0.99	N	0.00	Y	N
		(0.021)	(0.003)	(0.023)						
Peru	POLS	-1.290	0.020	0.335	0.00	0.00	Y	0.99	N	N
		(0.019)	(0.004)	(0.023)						
	CRE	-1.327	0.045	-0.067	0.00	0.00	Y	0.00	Y	Y
		(0.019)	(0.003)	(0.023)						
	FEIS	-1.244	0.046	-0.130	0.00	0.00	Y	0.00	Y	Y
		(0.022)	(0.003)	(0.023)						
Vietnam	POLS	-1.128	-0.005	0.347	0.00	0.91	N	0.99	N	N
		(0.019)	(0.004)	(0.023)						
	CRE	-1.130	0.003	0.024	0.00	0.13	N	0.85	N	N
		(0.019)	(0.003)	(0.023)						
	FEIS	-1.124	0.020	0.036	0.00	0.00	Y	0.94	N	N
		(0.023)	(0.002)	(0.023)						

Note: the first three columns report parameter estimates (and cluster-robust standard errors, in parentheses) from alternative estimators (POLS, CRE and FEIS); all estimates include a set of time-varying controls; columns denoted ‘H0’ report the probability associated with individual tests of the one-sided null hypotheses associated with each parameter (see Table 2); columns denoted ‘?’ report the conclusion of (composite) hypotheses regarding specific forms of catch-up growth; probability estimates larger than 0.99 have been rounded down to the latter value.
Source: own estimates.

Based on the parameter estimates, we test the catch-up hypotheses. Since these are built up from individual hypothesis tests, against each parameter we report the probability associated with the relevant one-sided null (denoted H0). To determine the specific combination of catch-up possibilities revealed by the procedure, we then combine the individual probabilities using the maxP methodology (Section 3). These results are summarized in the '?' columns, where a N(o) suggests we cannot reject the null hypothesis of no catch-up at the 5% confidence level; and a Y(es) indicates the presence of statistical evidence for a given form of catch-up. Given the chosen direction of the null hypotheses, we do not test for the presence of growth *divergence*.

Five main findings stand out. First, looking across the different estimators, the qualitative conclusions are broadly similar. However, parameter estimates are not identical within each country. In some cases, the estimates switch signs depending on the estimator chosen – e.g., mean growth velocity in Ethiopia is negative and significant under the POLS estimator (-0.016), but positive (0.003) and insignificant under the FEIS estimator. Furthermore, in all estimates, the FEIS estimates for β_0 are numerically larger than those taken from either the POLS or CRE models. Since the FEIS estimator imposes fewer assumptions, it is expected to be consistent under a wider range of conditions. In other words, the POLS and CRE estimators appear to yield downward-biased estimates of mean growth velocity, perhaps driven by unobserved individual heterogeneity. Additionally, the FEIS model displays the best goodness-of-fit. As reported in Appendix Table A3, the regression root mean square error (RMSE) and AIC criterion substantially improve (decline) as we move from the POLS to the CRE estimator; and in most cases they improve again when moving from the CRE to the FEIS estimator. The same table shows that estimation speed is much faster under the FEIS estimator compared to the CRE model, by a factor of around 100. Even more critically, Hausman tests comparing the CRE and FEIS results consistently reject the hypothesis that the estimates for the common parameters (namely, γ) are equal under both estimators. Taken together, these results suggest the FEIS approach is to be preferred. Thus, unless otherwise indicated, we focus on these results in our interpretation and henceforth.

Second, as expected given the purpose of the YL initiative is to study development trajectories among generally disadvantaged groups, the estimate for standardized mean stature ($\hat{\alpha}_0$) is below that of the reference population in all four countries. Consequently, a necessary precondition for between-population catch-up growth is fulfilled and the null hypothesis $H0_{\alpha_0} : \alpha_0 \geq 0$ is consistently rejected. Third, in keeping with the long difference results in Table 3(b), the pace of mean growth relative to the reference population ($\hat{\beta}_0$) differs across the four countries. In Peru and Vietnam, mean height z-scores in the sample are growing faster than in the reference population, implying we can reject the null hypothesis that $\beta_0 \leq 0$. In Ethiopia and India, by contrast, we find no evidence of above-normal growth rates. Thus, according to the FEIS results, between-group catch-up has taken place in Peru and Vietnam, but not elsewhere.

Fourth, evidence for within-group catch-up is also mixed. While the estimates for $\rho_{\alpha\beta}$ generally

observed across all four rounds. By construction, this procedure does not affect the correlation coefficient.

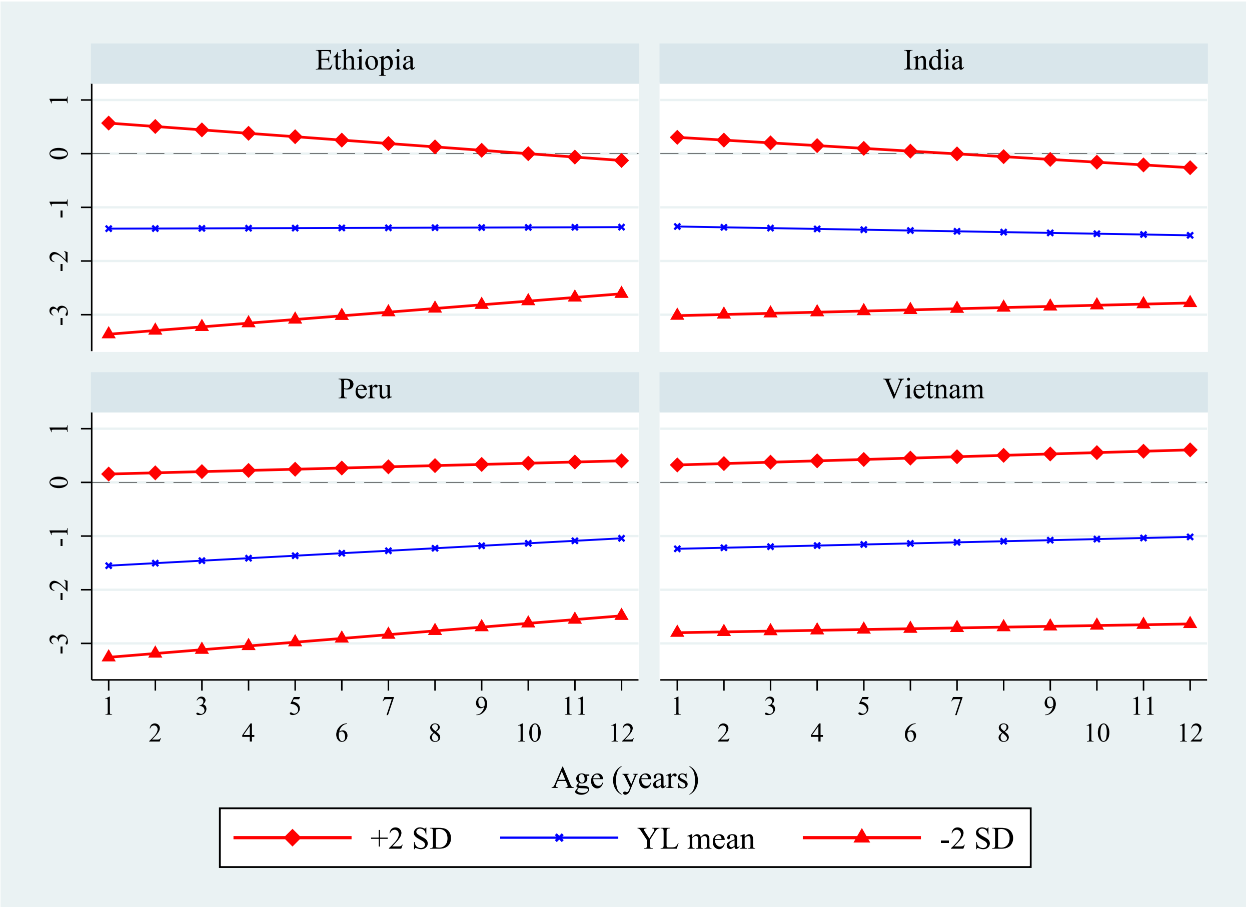
are closer to zero under the FEIS estimator as compared to at least one of the other estimators, the FEIS correlation coefficients are negative and significant in all locations other than Vietnam. It follows there is somewhat more consistent evidence of within-group catch-up of HAZ scores in the YL countries (i.e. in three of the four countries) than there is evidence of between-group catch-up (in two of the four countries). Taking the two forms of catch-up together, as tested in the last column of the table, only Peru displays evidence of *both* between- and within-group catch-up simultaneously.

Finally, while the above tests concern the presence of catch-up growth, they do not speak to its magnitude. Looking more closely at the size of the point estimates, we note the pace of between- and within-group catch-up is (at best) moderate and incomplete. For instance, while Peru displays the largest point estimate for β_0, it would take around 27 years (on the same trajectory) for the mean HAZ score in this sample to converge to zero. Figure 1 combines the preferred FEIS results for each country and plots simulated HAZ score trajectories by age. Alongside the trend of the YL sample mean, we plot the expected HAZ trends at both plus and minus two standard deviations of the estimated child-specific fixed-effects distribution (i.e., $\alpha_0 \pm 2\sigma_\alpha$), taking into account the estimated correlation with the slope effect. The results visually confirm the slow pace of between-group catch-up – i.e., over the 11 years spanned by the data, the mean (predicted) HAZ for Peru and Vietnam has increased by less than 0.5 of a standard deviation. Similarly, for those countries displaying within-country catch-up, this also proceeds slowly. The estimate for $\rho_{\alpha\beta}$ is largest in the case of Ethiopia. But even for this sample, the distance between the trends located at plus/minus two standard deviations declines by approximately 30% from ages 1 to 12; and these two trends would take about 30 years (on the same trajectories) to converge. These additional results underscore the merit of applying a latent growth model framework; but they also highlight the limited magnitude of catch-up in HAZ scores among the four YL countries.

4.3 Sensitivity analysis

The previous section focussed on results for HAZ scores only. While these scores are widely used to investigate catch-up growth, they are not the only viable metric. As noted above (Sections 2 and 4.1), previous studies use different transformations of height and BMI; and these may yield different conclusions about changes over time (also see Le and Behrman 2017, Zhang et al. 2016). To give one example, changes in externally standardized outcomes (e.g., HAZ scores) will reflect some combination of differences in heights *and* changes to the shape of the reference distribution (Cameron et al. 2005). Leroy et al. (2015) argue that changes in HAZ scores may provide overly optimistic evidence about growth catch-up because the dispersion of heights tends to increase with age – i.e., even if the mean absolute height difference between a sample and the median of a reference population remains constant, the corresponding z-score difference will tend to fall. As a result, the authors propose catch-up should be evaluated using absolute differences in height-for-age

Figure 1: Estimated HAZ score trajectories, by country



Note: trends are derived from the FEIS estimates reported in Table 4; the blue line (YL mean) is the estimate of the mean HAZ score for the sample; the two red lines are estimated at ± two standard deviations (SDs) of the estimated individual fixed-effects distributions; dashed horizontal line at $y = 0$ is the reference median. Source: own estimates.

(HAD) between the sample and a healthy reference population median. Based on this metric, the authors find evidence for growth faltering (deterioration in relative heights) across many low income countries, including those surveyed by the Young Lives project.

In light of these concerns, we investigate the extent to which the conclusions of Section 4.2 are sustained when other anthropometric outcomes are employed. To do so, we re-run the same analysis but now use a range of height-for-age transformations. The results are reported in Table 5, which replicates the format of Table 4 in summary form (dropping the probability values) and focussing uniquely on estimates from the FEIS model.[20] In line with the discussion around Table 3(b), as well as Figures A1–A3, different height transformations do yield substantively different conclusions. A key finding is that the HAD metric yields a much less positive view of catch-up. Based on this metric, there is no evidence of any form of catch-up in any country over time. Moreover, there appears to be material within-group *divergence* – estimates for $\rho_{\alpha\beta}$ in Table 5 are at least 39%

[20] The first set of results shown in Table 5 repeats those of Table 4. A complete set of estimates, based on other estimators, is available on request.

for the HAD metric. As already argued, this finding may reflect the arithmetic properties of the HAD measure, by which initial differences grow in absolute magnitude over time. The HAR metric corrects for this problem, but remains unaffected by changes in the distribution of the reference population. These results reported in Table 5 for the height ratio are generally less pessimistic than those based on absolute differences. Nonetheless, the HAR estimates continue to provide no evidence of within-group catch-up, and between-group catch-up only obtains for the Peruvian sample.

We also investigate catch-up growth based on various measures of BMI. These are shown in Table 6, which also only uses the FEIS estimator and summarizes results for various transformations. Here, the findings appear less sensitive to the particular metric chosen. There is no evidence of between-group catch-up for the continuous variables (BAZ, BAD, BAR); and evidence of within-group convergence is limited to Ethiopia (BAZ and BAR metrics) and Peru (BAZ only). The greater consistency of findings in this case may well reflect the properties of BMI. Since both height and weight are growing in a healthy reference population, changes in (median) BMI are expected to be fairly stable between ages one and 12. Also, the distribution of body mass shows rather less variation (changes in variance) across different age groups compared to the distribution of heights. Thus, for the BMI measure changes in dispersion are likely to play a smaller role, in turn explaining the broad similarity of findings across the continuous transformations.

The same analysis can be applied to binary indicators, in which case the outcome is a dummy variable and the specification becomes a linear probability model. Corresponding results are shown in the bottom portions of Tables 5 and 6), where the binary indicators take a value of one if a child is within a healthy group (this construction ensures the hypothesis tests can be applied in the same way as before). Three points stand out. First, consistent with the HAZ results, there is clear evidence of both between- and within-group catch-up from stunting (however defined) in the Peruvian sample. Second, all other countries show some sign of a reduction in the prevalence of stunting – i.e., there is some between-group catch-up. But this depends on the specific metric chosen. In Ethiopia, there is a reduction in both the prevalence of stunting (HAZ < -2) and severe stunting (HAZ < -3). In India, we only see a fall in severe stunting; and in Vietnam we observe a reduction in stunting, but not severe stunting. Indeed, all countries other than Vietnam show evidence of between- and within-group catch-up from severe stunting. However, this may well reflect the lower prevalence of severe stunting in Vietnam to begin with (see Table 3). Third, there is limited evidence of catch-up from wasting. Indeed, none of the countries provides evidence of either between- or within-group catch-up from very low BMI. Children from the Peru sample show some increase in the share of children with a healthy BMI. However, this is likely to be driven by a fall in those with a very high initial BMI (obese).

As a final exercise, Table 7 pulls together the results from the various hypothesis tests examined above (for the FEIS estimator). Similar to Table 2, it counts the number of test results corresponding to each combination of within and between catch-up. Part (a) of the table refers to the five height

Table 5: Estimates of catch-up growth in height-for-age outcomes (conditional model)

	Between			Within		Both
	α_0	β_0	?	$\rho_{\alpha\beta}$	?	?
Externally standardized difference (HAZ):						
Ethiopia	-1.38	0.00	N	-0.30	Y	N
India	-1.44	-0.01	N	-0.19	Y	N
Peru	-1.24	0.05	Y	-0.13	Y	Y
Vietnam	-1.12	0.02	Y	0.04	N	N
Ratio to reference median (HAR$\times$100):						
Ethiopia	-5.76	-0.13	N	-0.01	N	N
India	-6.07	-0.22	N	0.14	N	N
Peru	-5.36	0.10	Y	0.02	N	N
Vietnam	-4.71	-0.01	N	0.37	N	N
Difference to reference median (HAD):						
Ethiopia	-6.89	-0.52	N	0.53	N	N
India	-7.29	-0.61	N	0.64	N	N
Peru	-6.54	-0.18	N	0.39	N	N
Vietnam	-5.60	-0.28	N	0.76	N	N
Not stunted (HAZ > -2):						
Ethiopia	-26.75	0.56	Y	0.06	N	N
India	-27.85	-0.12	N	0.12	N	N
Peru	-22.22	1.06	Y	-0.30	Y	Y
Vietnam	-18.45	0.39	Y	0.04	N	N
Not severely stunted (HAZ > -3):						
Ethiopia	-9.02	0.76	Y	-0.28	Y	Y
India	-6.59	0.41	Y	-0.11	Y	Y
Peru	-5.33	0.59	Y	-0.54	Y	Y
Vietnam	-3.89	0.01	N	0.05	N	N

Note: columns $\alpha_0, \beta_0, \rho_{\alpha\beta}$ report parameter estimates based on the FEIS estimator; all specifications include a set of time-varying controls; columns denoted '?' report the decision (N = no; Y = yes) of whether we can reject the (composite) null hypotheses regarding the presence of specific forms of catch-up growth.
Source: own estimates.

Table 6: Estimates of catch-up growth in BMI-for-age outcomes (conditional model)

	Between			Within		Both
	α_0	β_0	?	$\rho_{\alpha\beta}$	?	?
Externally standardized difference (BAZ):						
Ethiopia	-1.08	-0.13	N	-0.09	Y	N
India	-1.24	-0.02	N	0.23	N	N
Peru	0.59	-0.02	N	-0.10	Y	N
Vietnam	-0.51	-0.03	N	0.37	N	N
Ratio to reference median (BAR$\times$100):						
Ethiopia	-8.98	-1.33	N	-0.05	Y	N
India	-9.67	-0.23	N	0.68	N	N
Peru	8.33	0.34	N	0.30	N	N
Vietnam	-3.45	-0.18	N	0.62	N	N
Difference to reference median (BAD):						
Ethiopia	-1.50	-0.25	N	0.00	N	N
India	-1.58	-0.05	N	0.71	N	N
Peru	1.37	0.08	N	0.38	N	N
Vietnam	-0.57	-0.04	N	0.66	N	N
BMI-for-age is not low (> 5th percentile):						
Ethiopia	-25.80	-4.69	N	0.36	N	N
India	-30.05	-1.26	N	0.14	N	N
Peru	3.42	-0.08	N	-0.03	N	N
Vietnam	-9.36	-1.66	N	0.41	N	N
BMI-for-age is healthy (5th – 85th percentile):						
Ethiopia	-17.91	-3.53	N	0.18	N	N
India	-19.04	-1.50	N	0.10	N	N
Peru	-12.29	0.45	Y	-0.22	Y	Y
Vietnam	-4.13	-2.01	N	0.32	N	N

Note: columns $\alpha_0, \beta_0, \rho_{\alpha\beta}$ report parameter estimates based on the FEIS estimator; all specifications include a set of time-varying controls; columns denoted '?' report the decision (N = no; Y = yes) of whether we can reject the (composite) null hypotheses regarding the presence of specific forms of catch-up growth.
Source: own estimates.

measures; and part (b) of the table refers to the five BMI measures. It shows that while there is certainly no overwhelming evidence for catch-up of either form, around half of all tests (models) reveal at least some form of catch-up in stature. In contrast, 4 in every 5 tests for the BMI outcomes reveal no form of catch-up whatsoever. The disconnect between the findings for these two (related) outcomes is striking and merits future research. The table also provides insights regarding the relationship between the two forms of catch-up. Fisher's exact test of their relationship suggests we cannot reject the null hypothesis of independence. This supports the contention that the different forms of catch-up are distinct and merit separate attention.

To conclude this subsection, it is evident that estimates of both forms of catch-up are sensitive to the choice of metric. Indeed, how outcomes are expressed often seems rather more fundamental than the choice of regression estimator, especially if we exclude a naïve POLS approach (i.e., the CRE and FEIS estimators yield qualitatively similar conclusions). This concern applies particularly to outcomes associated with height. Here, conclusions regarding catch-up based on a ratio transformation tend to fall between the more optimistic results deriving from z-score transformations (e.g., HAZ and binary measures of stunting) and the pessimistic results that emerge when we use absolute gaps (HAD). Differences between outcome metrics reflect a range of factors, including (*inter alia*): the effect of changes in the shape of the reference distribution; alternative ways in which height and weight are combined; how initial gaps compound over time; and the focus on particular aspects of the outcome distribution. Since the construction choices underlying specific outcomes often correspond to different substantive research interests, there is unlikely to be a universal 'best' metric or transformation. Nonetheless, in the absence of a strong *a priori* justification, the evidence presented here suggests researchers should be wary of uniquely relying on any single outcome indicator to make general claims about catch-up growth.

4.4 Growth modifiers

Thus far, our analysis has concentrated on the broad direction and magnitude of catch-up growth (if any). However, we have not considered which factors might account for variation in individual child growth trajectories within each sample. In the existing literature, one approach to doing so focuses on simplified versions of equation (1). Specifically, variation in the estimated residuals is analyzed: $\hat{\varepsilon}_{it} = y_{it} - \mathrm{E}(y_{it} \mid y_{i,t-1}, x_{is})$, including their association with later outcomes such as cognitive achievement (e.g., Crookston et al. 2013, Gandhi et al. 2011, Kuklina et al. 2006, Schott et al. 2013). Another approach, based on direct models of growth trajectories, is to consider potential heterogeneity in the β coefficient in equation (6a), which can be examined by including interaction terms between age and relevant explanatory variables (e.g., Rieger and Trommlerová 2016, Rubio-Codina et al. 2015).

Our extension borrows on elements of both approaches. As discussed by Cole et al. (2010), individual growth curves can be characterized by a small number of parameters. In our case, given

Table 7: Frequency tabulation of hypotheses tests, by outcome type

(a) *Height outcomes*:

Within?	Between? Yes	No	Total
Yes	5	2	7
No	4	9	13
Total	9	11	20

(b) *BMI outcomes*:

Within?	Between? Yes	No	Total
Yes	1	3	4
No	0	16	16
Total	1	19	20

Note: table reports counts of between-group and within-group null hypothesis test results, derived from Tables 5 and 6, covering all countries.
Source: own estimates.

we assume a linear specification, the individual growth trajectories are simply defined by child mean size: $\tilde{\alpha}_i = \hat{\alpha}_0 + \hat{\alpha}_i$; and linear growth velocity: $\tilde{\beta}_i = \hat{\beta}_0 + \hat{\beta}_i$. So, rather than considering variation in β, alone, we use the FEIS estimates for both variables and treat them as dependent variables. Concretely, this stage of analysis is based on regressions of the following form where children are the unit of analysis (in cross-section):

$$\tilde{\alpha}_i = \delta_0 + \bar{x}_i'\delta_1 + \delta_{2c} + \eta_i \tag{7a}$$

$$\tilde{\beta}_i = \theta_0 + \bar{x}_i'\theta_1 + \theta_{2c} + \nu_i \tag{7b}$$

Estimates for δ_1 indicate the impact of the mean values of the vector x, defined as before, on average child size. Estimates for θ_1 give the corresponding impacts on growth velocity, indicating how the permanent elements of x modify the pace of child development. Recall the estimated latent variables absorb the effects of all time-invariant child characteristics. So, this analysis seeks to quantify the extent to which these are accounted for by observed child characteristics. In addition, community-level fixed effects can be added to the regression to capture how differential access to public services and other features of the broad local environment (outside the family) systematically affect child growth trajectories. These are captured by the parameters δ_{2c}, θ_{2c}, where c is an index for the community.[21]

This kind of second stage analysis has been applied in other contexts. Acemoglu et al. (2009), for

[21] When treated as fixed effects, we cannot identify which specific features of the local environment (community) alter child growth trajectories. Nonetheless, as we show, this approach helps quantify the overall explanatory contribution of variables at this level.

example, investigate the modernization hypothesis by regressing a set of country-specific fixed effects against time-invariant country characteristics; and Brand and Davis (2011) use a similar method to assess heterogeneity in the effects of a college education on fertility. One challenge in using this two-stage approach is that α_i and β_i are estimated with error. To address this, we follow the advice of Lewis and Linzer (2005) who suggest that, in small samples, unweighted OLS regressions combined with Efron or bootstrap consistent standard errors are generally reliable when the dependent variables are themselves estimates.

For the remainder of the paper we focus on the height and BMI ratio outcomes. We do so as the ratio transformation represents the lesser of two extremes – the ratio of outcomes in the sample to that of the median in the reference distribution is not affected by (independent) changes in the variance of the reference distribution, nor does it mechanically compound initial differences. Tables 8 to 11 summarize our results for the HAR and BAR outcomes, covering each country separately. Appendix Tables A4 to A7 report the same analysis for the corresponding z-scores (HAZ and BAZ). In each table, the first group of the columns (Ia-Ie) refer to the height outcomes; and the final set of columns refer to body mass outcomes. Regression results for mean child size are reported in columns Ia, Ib, IIa and IIb where the first in each set excludes community fixed effects. Columns Ic, Id, IIc and IId report the same regressions for growth velocity. Lastly, columns Ie and IIe add the estimates for child size ($\tilde{\alpha}_i$) to the velocity regression, reflecting the potential correlation between these two variables (as captured by $\rho_{\alpha\beta}$; see above). This extension is informative because it helps differentiate between the contributions of observed and unobserved individual determinants of growth velocity.

We highlight four main findings from these results. First, a number of the coefficient estimates are consistent with previous literature. In particular, household wealth displays a significant positive relationship with both size and velocity across most countries.[22] Similarly, the highest school grade attained by the mother, here transformed into standard deviation units, shows a significant positive association with predicted child size (especially height), but a less systematic relation with growth velocity. Overall, the implication is that children from more advantaged households show enhanced average development outcomes and would appear to extend their advantage over time relative to other children in the sample. However, these advantages generally do not imply that children from (relatively) wealthy families catch up toward the median of the external reference population. This is indicated in the final row of the table, which calculates the predicted size and growth velocity among children from households that are two standard deviations above the country mean wealth index level. Across all the regressions for height, the adjusted estimates of mean stature remain in the negative domain; and only in Peru and Vietnam is mean growth velocity in the positive domain for the more wealthy households. In this sense, between-group catch-up remains elusive in both Ethiopia and India, even for children from more advantaged households in the sample.

[22] The wealth index is constructed as the standardized first principal component of three separate indexes, which are provided in the YL datasets. These pertain to ownership of consumer durables, housing quality, and access to basic amenities within the household (e.g., clean water).

Table 8: Size and velocity regressions, Ethiopia

	Height ratios					BMI ratios				
	Size Ia	Size Ib	Velocity Ic	Velocity Id	Velocity Ie	Size IIa	Size IIb	Velocity IIc	Velocity IId	Velocity IIe
Mean	-5.781***	-5.774***	-0.129***	-0.129***	-0.129***	-8.988***	-8.998***	-1.328***	-1.326***	-1.327***
	(0.074)	(0.074)	(0.009)	(0.010)	(0.008)	(0.107)	(0.121)	(0.021)	(0.024)	(0.022)
Female	0.375**	0.412***	-0.027	-0.020	-0.019	-1.429***	-1.475***	-0.111***	-0.126***	-0.150***
	(0.155)	(0.148)	(0.019)	(0.018)	(0.019)	(0.212)	(0.247)	(0.040)	(0.037)	(0.040)
Urban	-0.265	1.096	0.014	0.111	0.115	-0.102	1.330	0.105	0.324	0.345
	(0.258)	(0.832)	(0.031)	(0.146)	(0.136)	(0.368)	(1.681)	(0.067)	(0.335)	(0.294)
Not first born	-0.032	-0.122	-0.018	-0.002	-0.002	-0.535	-0.273	0.187***	0.177***	0.173***
	(0.217)	(0.190)	(0.028)	(0.026)	(0.032)	(0.354)	(0.390)	(0.063)	(0.055)	(0.056)
Majority ELG	0.358*	0.257	-0.038	0.027	0.028	-1.236***	-0.179	-0.260***	-0.084	-0.086
	(0.183)	(0.310)	(0.023)	(0.035)	(0.033)	(0.301)	(0.433)	(0.054)	(0.081)	(0.075)
Wealth index	1.143***	1.147***	-0.005	-0.024	-0.020	-0.089	-0.569**	-0.116***	-0.083**	-0.092**
	(0.130)	(0.130)	(0.018)	(0.019)	(0.017)	(0.196)	(0.244)	(0.036)	(0.038)	(0.039)
Mother's edu.	0.228*	0.186	-0.038**	-0.029*	-0.028*	0.168	-0.196	0.000	0.025	0.022
	(0.117)	(0.121)	(0.016)	(0.016)	(0.017)	(0.204)	(0.231)	(0.038)	(0.037)	(0.037)
Mother's age	0.004	0.006	-0.004**	-0.004**	-0.004**	0.020	0.017	0.004	0.002	0.002
	(0.014)	(0.015)	(0.002)	(0.002)	(0.002)	(0.027)	(0.023)	(0.004)	(0.004)	(0.004)
Sibs 0-5	-0.495**	-0.492**	-0.104***	-0.057**	-0.059**	-0.647**	-0.399	-0.027	-0.198***	-0.204***
	(0.212)	(0.195)	(0.025)	(0.026)	(0.025)	(0.320)	(0.326)	(0.056)	(0.050)	(0.050)
Sibs 6-12	-0.063	-0.054	-0.062**	-0.051*	-0.052*	-0.700**	-0.796**	0.028	0.025	0.012
	(0.178)	(0.196)	(0.029)	(0.027)	(0.027)	(0.327)	(0.332)	(0.059)	(0.048)	(0.053)
Child FE (std.)	-	-	-	-	-0.013	-	-	-	-	-0.086***
					(0.011)					(0.029)
Obs.	1,845	1,845	1,845	1,845	1,845	1,846	1,846	1,846	1,846	1,846
RMSE	3.10	3.03	0.39	0.38	0.38	5.23	5.01	0.91	0.82	0.81
Adj. R^2	0.13	0.17	0.02	0.09	0.09	0.05	0.13	0.04	0.22	0.23
WI$\times$2 + Mean	-3.50	-3.48	-0.14	-0.18	-0.17	-9.17	-10.14	-1.56	-1.49	-1.51
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.
Source: own estimates.

Table 9: Size and velocity regressions, India

	Height ratios					BMI ratios				
	Size Ia	Size Ib	Velocity Ic	Velocity Id	Velocity Ie	Size IIa	Size IIb	Velocity IIc	Velocity IId	Velocity IIe
Mean	-6.061***	-6.061***	-0.222***	-0.222***	-0.222***	-9.682***	-9.681***	-0.224***	-0.223***	-0.222***
	(0.065)	(0.066)	(0.009)	(0.009)	(0.008)	(0.178)	(0.139)	(0.036)	(0.039)	(0.033)
Female	0.272**	0.195	0.010	0.011	0.008	-0.087	-0.138	0.061	0.044	0.067
	(0.132)	(0.131)	(0.017)	(0.015)	(0.017)	(0.272)	(0.257)	(0.069)	(0.073)	(0.059)
Urban	-0.163	-0.112	0.065***	-0.024	-0.023	0.730	-0.843	0.123	-0.340	-0.200
	(0.203)	(0.674)	(0.024)	(0.087)	(0.094)	(0.500)	(1.569)	(0.091)	(0.315)	(0.298)
Not first born	-0.476***	-0.390**	-0.048**	-0.047**	-0.042*	-0.249	-0.326	-0.048	-0.071	-0.017
	(0.165)	(0.164)	(0.022)	(0.022)	(0.022)	(0.333)	(0.385)	(0.069)	(0.090)	(0.059)
Majority ELG	-0.255**	-0.133	0.000	-0.015	-0.014	-1.010***	-0.540	-0.284***	-0.169***	-0.079
	(0.130)	(0.140)	(0.017)	(0.017)	(0.014)	(0.311)	(0.329)	(0.072)	(0.057)	(0.058)
Wealth index	0.662***	0.642***	0.047***	0.060***	0.052***	1.458***	1.312***	0.088	0.072	-0.146
	(0.101)	(0.108)	(0.011)	(0.013)	(0.013)	(0.250)	(0.245)	(0.067)	(0.065)	(0.126)
Mother's edu.	0.223**	0.170*	-0.002	-0.008	-0.010	0.487	0.487	0.144	0.139	0.059
	(0.107)	(0.099)	(0.011)	(0.012)	(0.013)	(0.366)	(0.319)	(0.101)	(0.099)	(0.038)
Mother's age	0.046***	0.050***	-0.006***	-0.006**	-0.006***	-0.006	-0.000	-0.002	-0.002	-0.002
	(0.015)	(0.018)	(0.002)	(0.002)	(0.002)	(0.036)	(0.033)	(0.008)	(0.008)	(0.007)
Sibs 0-5	-0.038	0.029	0.048*	0.051**	0.051**	-0.068	-0.116	-0.102	-0.148	-0.128
	(0.223)	(0.217)	(0.027)	(0.025)	(0.025)	(0.587)	(0.538)	(0.108)	(0.146)	(0.104)
Sibs 6-12	-0.015	0.090	0.019	0.003	0.002	-0.491	-0.189	-0.132**	-0.092	-0.061
	(0.195)	(0.196)	(0.023)	(0.023)	(0.019)	(0.425)	(0.380)	(0.063)	(0.065)	(0.070)
Child FE (std.)	-	-	-	-	0.040***	-	-	-	-	1.260**
					(0.009)					(0.511)
Obs.	1,913	1,913	1,913	1,913	1,913	1,913	1,913	1,913	1,913	1,913
RMSE	2.90	2.86	0.35	0.33	0.33	7.30	7.28	1.79	1.79	1.32
Adj. R^2	0.08	0.11	0.05	0.13	0.14	0.07	0.08	0.02	0.03	0.47
WI×2 + Mean	-4.74	-4.78	-0.13	-0.10	-0.12	-6.77	-7.06	-0.05	-0.08	-0.51
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.
Source: own estimates.

Table 10: Size and velocity regressions, Peru

	Height ratios					BMI ratios				
	Size Ia	Size Ib	Velocity Ic	Velocity Id	Velocity Ie	Size IIa	Size IIb	Velocity IIc	Velocity IId	Velocity IIe
Mean	-5.354***	-5.351***	0.098***	0.097***	0.097***	8.332***	8.336***	0.343***	0.343***	0.343***
	(0.079)	(0.071)	(0.009)	(0.009)	(0.010)	(0.181)	(0.209)	(0.036)	(0.037)	(0.034)
Female	-0.094	-0.125	0.038**	0.031	0.031	-1.335***	-1.333***	-0.092	-0.089	-0.038
	(0.125)	(0.145)	(0.019)	(0.019)	(0.021)	(0.391)	(0.376)	(0.070)	(0.086)	(0.072)
Urban	0.367*	-0.211	-0.022	0.011	0.011	-0.982*	-1.421**	0.164	0.050	0.106
	(0.212)	(0.257)	(0.030)	(0.038)	(0.032)	(0.558)	(0.657)	(0.104)	(0.130)	(0.133)
Not first born	0.045	-0.063	-0.033	-0.035	-0.035	-0.437	-0.715	-0.003	-0.021	0.007
	(0.155)	(0.189)	(0.028)	(0.025)	(0.025)	(0.555)	(0.516)	(0.100)	(0.087)	(0.093)
Majority ELG	0.355*	0.069	-0.043	-0.036	-0.036	-0.546	-0.099	0.117	0.153	0.157
	(0.187)	(0.167)	(0.027)	(0.032)	(0.027)	(0.572)	(0.532)	(0.105)	(0.129)	(0.121)
Wealth index	0.913***	0.824***	0.022*	0.048***	0.050***	2.107***	1.619***	0.362***	0.248***	0.185***
	(0.084)	(0.100)	(0.012)	(0.018)	(0.015)	(0.248)	(0.279)	(0.045)	(0.061)	(0.056)
Mother's edu.	0.530***	0.471***	-0.010	-0.008	-0.007	0.435	0.454*	-0.012	0.001	-0.017
	(0.091)	(0.084)	(0.013)	(0.014)	(0.014)	(0.302)	(0.269)	(0.054)	(0.061)	(0.059)
Mother's age	0.019*	0.024**	-0.004**	-0.004**	-0.004**	0.040	0.046	-0.013**	-0.011*	-0.013*
	(0.011)	(0.011)	(0.002)	(0.002)	(0.002)	(0.037)	(0.032)	(0.006)	(0.006)	(0.006)
Sibs 0-5	0.076	0.087	-0.076***	-0.076***	-0.076***	0.003	-0.140	-0.270***	-0.289***	-0.283***
	(0.167)	(0.172)	(0.024)	(0.027)	(0.023)	(0.508)	(0.450)	(0.102)	(0.105)	(0.093)
Sibs 6-12	-0.669***	-0.496***	0.023	-0.002	-0.003	-0.523	0.063	0.107	0.154	0.151
	(0.192)	(0.184)	(0.035)	(0.027)	(0.031)	(0.512)	(0.457)	(0.105)	(0.104)	(0.116)
Child FE (std.)	-	-	-	-	-0.007	-	-	-	-	0.354***
					(0.015)					(0.057)
Obs.	1,916	1,916	1,916	1,916	1,916	1,921	1,921	1,921	1,921	1,921
RMSE	2.85	2.78	0.42	0.41	0.41	8.67	8.28	1.63	1.62	1.59
Adj. R^2	0.31	0.35	0.02	0.05	0.05	0.10	0.17	0.11	0.13	0.16
WI$\times$2 + Mean	-3.53	-3.70	0.14	0.19	0.20	12.55	11.57	1.07	0.84	0.71
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.
Source: own estimates.

Table 11: Size and velocity regressions, Vietnam

	Height ratios					BMI ratios				
	Size Ia	Size Ib	Velocity Ic	Velocity Id	Velocity Ie	Size IIa	Size IIb	Velocity IIc	Velocity IId	Velocity IIe
Mean	-4.709***	-4.709***	-0.012*	-0.012*	-0.012*	-3.433***	-3.429***	-0.181***	-0.181***	-0.182***
	(0.060)	(0.071)	(0.007)	(0.006)	(0.007)	(0.157)	(0.183)	(0.026)	(0.026)	(0.023)
Female	0.187	0.139	0.005	0.006	0.001	-1.764***	-1.720***	-0.159***	-0.162***	-0.016
	(0.123)	(0.135)	(0.014)	(0.013)	(0.013)	(0.348)	(0.343)	(0.049)	(0.048)	(0.039)
Urban	1.313***	0.843	-0.001	0.013	-0.016	4.646***	6.147***	0.177**	0.195	-0.327
	(0.213)	(0.653)	(0.023)	(0.087)	(0.090)	(0.602)	(1.826)	(0.076)	(0.279)	(0.235)
Not first born	-0.451**	-0.383**	-0.019	-0.017	-0.004	-1.131**	-1.125**	-0.154**	-0.138**	-0.043
	(0.182)	(0.163)	(0.019)	(0.020)	(0.018)	(0.478)	(0.477)	(0.070)	(0.063)	(0.061)
Majority ELG	1.272***	0.294	0.005	-0.050	-0.060	-1.473***	2.536**	0.013	-0.285*	-0.500***
	(0.183)	(0.318)	(0.019)	(0.053)	(0.052)	(0.452)	(1.053)	(0.062)	(0.171)	(0.176)
Wealth index	0.648***	0.893***	0.054***	0.057***	0.027**	0.948***	1.550***	0.226***	0.291***	0.159***
	(0.107)	(0.108)	(0.012)	(0.014)	(0.012)	(0.266)	(0.311)	(0.040)	(0.045)	(0.036)
Mother's edu.	0.567***	0.580***	0.023**	0.014	-0.006	0.815***	0.808**	-0.010	-0.002	-0.071**
	(0.108)	(0.104)	(0.010)	(0.012)	(0.013)	(0.313)	(0.316)	(0.044)	(0.043)	(0.036)
Mother's age	-0.017	-0.024	-0.006***	-0.005***	-0.004**	0.040	-0.001	0.012**	0.009*	0.009**
	(0.014)	(0.015)	(0.001)	(0.001)	(0.002)	(0.037)	(0.034)	(0.005)	(0.005)	(0.004)
Sibs 0-5	-0.895***	-0.459	-0.026	-0.013	0.003	0.528	0.324	-0.136	-0.091	-0.119
	(0.284)	(0.293)	(0.031)	(0.034)	(0.030)	(0.687)	(0.729)	(0.100)	(0.088)	(0.097)
Sibs 6-12	-0.434	0.022	-0.055**	-0.046*	-0.046**	-1.097	0.212	-0.230***	-0.104	-0.122
	(0.267)	(0.289)	(0.027)	(0.026)	(0.023)	(0.684)	(0.638)	(0.087)	(0.090)	(0.087)
Child FE (std.)	-	-	-	-	0.115***	-	-	-	-	0.684***
					(0.010)					(0.028)
Obs.	1,908	1,908	1,908	1,908	1,908	1,909	1,909	1,909	1,909	1,909
RMSE	2.85	2.74	0.30	0.30	0.28	7.47	7.19	1.07	1.04	0.85
Adj. R^2	0.29	0.35	0.07	0.10	0.19	0.14	0.20	0.10	0.14	0.44
WI×2 + Mean	-3.41	-2.92	0.10	0.10	0.04	-1.54	-0.33	0.27	0.40	0.14
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.
Source: own estimates.

The second insight, which follows directly, is the presence of notable differences between countries in the magnitude (and significance) of different explanatory factors. For instance, there is considerable heterogeneity in the relative trajectories of boys and girls within each sample. In both Ethiopia and Vietnam, girls show a larger negative average body mass difference to the gender-specific reference population, and also show more rapid divergence compared to boys over time. In India, by contrast, no systematic height or body mass differences are found between boys and girls, at least after controlling for community characteristics (Table 9). Nonetheless, a distinctive insight from India is the strong relationship between birth order and stature. In keeping with recent evidence due to Jayachandran and Pande (2017), first borns are both taller and grow faster than their younger siblings. Results for the Peruvian sample indicate no birth order effects; younger sibs in Ethiopia tend to gain body mass faster than first borns; but younger sibs in Vietnam show a lower BMI than their older sibs, and this gap grows over time.

Third, community characteristics play a material role in explaining differences in growth trajectories. This is seen in two ways. First, inclusion of the community fixed effects leads to material changes in the regression coefficient estimates in may cases. For instance, in the estimates that exclude community fixed effects, we find that children from the region's majority ethno-linguistic group display different growth trajectories to other children, especially in India and Vietnam. However, these differences become substantially moderated once community fixed effects are included. Second, the (adjusted) R^2 estimates, reported in the footer of the regression tables, all increase on inclusion of the community fixed effects. Indeed, the increase relative to the same regression without the fixed effects represents a lower bound on the variance share attributable to variables at this level. This varies across models, but is around 5 percentage points in most cases, which is material in relation to the total (adjusted) R^2.

Lastly, notwithstanding the above, we recognize that the included covariates explain a comparatively small share of the overall variation in the estimated latent variables. Again, this is shown by the (adjusted) R^2 estimates. For the mean size (height) ratio estimates, less than one-third of the variance can be accounted for across all the estimates. While some of this may be attributable to measurement error, it is reasonable to conclude that the set of available controls does not provide a comprehensive picture of differences in stature and body mass between children. The same general point extends to the velocity estimates. However, in a number of cases, the explanatory power of the regression increases significantly when the latent child fixed effect (size) is added. For instance, the unexplained component of growth in BMI differences in India falls from 98% to 53% once $\tilde{\alpha}_i$ is added. This indicates that unobserved determinants of child development, perhaps including genetic variation and other aspects of the home environment, are fundamental. Moreover, inclusion of $\tilde{\alpha}_i$ often alters the magnitude (and even sign) of some observed coefficients. This points out that results for the latter may be biased by omitted variables, even after controlling for unobserved community characteristics – e.g., wealth may be correlated with unobserved household conditions. Thus, we should be wary about drawing strong conclusions about the potential impact of specific policy interventions based on descriptive features of growth trajectories. Put differently, while these regressions are of descriptive interest, they are unlikely to yield precise causal estimates.

5 Conclusion

The aim of this study was to revisit how catch-up growth in children is defined and measured. Our motivation was that existing definitions tend to focus on specific dimensions of catch-up – namely either between-group or within-group convergence, but not both. We argued that existing studies also tend to focus on specific outcomes (e.g., HAZ) and that conventional dynamic panel estimates of catch-up are prone to bias. In contrast, we suggested that a latent growth framework is an attractive alternative. It avoids many of the technical problems with dynamic panel approaches and provides a unified setting in which to test for both between- and within-group catch-up. We outlined how the latent growth framework can be estimated in practice, proposing that a fixed-effects and individual slopes (FEIS) estimator imposes minimal assumptions relative to alternatives such as a correlated random effects (mixed effects) model. In addition, we showed how composite hypotheses regarding catch-up can be tested by combining probability values from the regression estimates.

In our application of this approach, we compared the performance of various latent growth estimators using four rounds of the Young Lives data collected in (sentinel sites) in Ethiopia, India, Peru and Vietnam. All of the samples come from relatively disadvantaged populations and, as expected, on average the sampled children display lower height and body mass than those from a healthy international reference population. On the basis of externally-standardized HAZ scores, our latent growth estimates revealed there has been statistically significant catch-up growth in Peru and Vietnam, but that the speed of catch-up has been slow. All countries other than Vietnam were also found to display some degree of within-group catch-up, but again this appears to have proceeded at a relatively moderate pace. Together, and as depicted in Figure 1, these results confirm the importance of conceptually and practically distinguishing between within- and between-group catch-up. Indeed, while the (predicted) distribution of HAZ scores for the sample in India is narrowing, the overall trend for the sample is declining. Only in Peru do we find evidence for both forms of catch-up simultaneously.

In terms of technical insights, we found that the FEIS estimator outperforms existing alternatives. Estimates based on a POLS or CRE estimator were found to be inconsistent (biased) and showed poorer goodness-of-fit. At the same time, conclusions about catch-up were found to be highly sensitive to the outcome metric chosen, especially in the case of height outcomes. As found elsewhere, evidence on catch-up based on HAZ scores tends to be more optimistic than evidence based on absolute or relative height differences. We argued that while there is unlikely to be a single 'best' metric, the choice of outcome metric is far from trivial and needs to be justified in relation to substantive research questions. Moreover, checking whether findings about catch-up hold across alternative metrics can be very helpful. Indeed, our analysis showed a clear disconnect in the evidence for catch-up based on metrics of stature vs BMI (see Table 7).

Lastly, we extended the analysis to show how the framework can be used to identify systematic drivers of heterogeneity in both size and growth velocity. This corresponded to a second analytical stage, in which the estimated latent variables (i.e., the FEIS intercept and slope fixed effects) were regressed on (time-invariant) individual characteristics such as gender, household wealth and the mother's level of education as well as community fixed effects. These results revealed significant diversity across the four countries. However, higher household wealth is often associated with more rapid growth, enabling children to extend initial advantages. Also, community fixed effects appear to play a material role in influencing growth trajectories. In sum, a latent growth framework combined with a flexible FEIS estimation procedure provides a rich and practical empirical basis, both to distinguish between alternative forms of catch-up and also to investigate factors affecting the heterogeneity in child growth patterns

References

Acemoglu, D., Johnson, S., Robinson, J.A. and Yared, P. (2009). Reevaluating the modernization hypothesis. *Journal of Monetary Economics*, 56(8):1043–1058.

Adair, L.S. (1999). Filipino children exhibit catch-up growth from age 2 to 12 years. *The Journal of Nutrition*, 129(6):1140–1148.

Almond, D., Currie, J. and Duque, V. (2017). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, forthcoming.

Anand, P., Roope, L. and Peichl, A. (2016). Wellbeing Evidence for the Assessment of Progress. IZA Discussion Papers 9840, Institute for the Study of Labor (IZA). URL https://ideas.repec.org/p/iza/izadps/dp9840.html.

Baltagi, B. (2008). *Econometric analysis of panel data*. Chicester: John Wiley & Sons.

Barnett, I., Ariana, P., Petrou, S., Penny, M.E., Le, T.D., Galab, S., Woldehanna, T., Escobal, J.A., Plugge, E. and Boyden, J. (2012). Cohort profile: the Young Lives study. *International Journal of Epidemiology*, 42(3):701–708.

Bell, A. and Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1):133–153.

Biesanz, J.C., Deeb-Sossa, N., Papadakis, A.A., Bollen, K.A. and Curran, P.J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods*, 9(1):30.

Black, M.M., Walker, S.P., Fernald, L.C., Andersen, C.T., DiGirolamo, A.M., Lu, C., McCoy, D.C., Fink, G., Shawar, Y.R., Shiffman, J. et al. (2017). Early childhood development coming of age: science through the life course. *The Lancet*, 389(10064):77–90.

Blance, A., Tu, Y.K. and Gilthorpe, M.S. (2005). A multilevel modelling solution to mathematical coupling. *Statistical Methods iIn Medical Research*, 14(6):553–565.

Bond, S.R. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, 1(2):141–162.

Bond, S., Nauges, C. and Windmeijer, F. (2005). Unit roots: identification and testing in micro panels. CeMMAP working papers CWP07/05, Centre for Microdata Methods and Practice, Institute for Fiscal Studies. URL https://ideas.repec.org/p/ifs/cemmap/07-05.html.

Brand, J.E. and Davis, D. (2011). The impact of college education on fertility: Evidence for heterogeneous effects. *Demography*, 48(3):863–887.

Cameron, N., Preece, M.A. and Cole, T.J. (2005). Catch-up growth or regression to the mean? Recovery from stunting revisited. *American Journal of Human Biology*, 17(4):412–417.

Chirwa, E.D., Griffiths, P.L., Maleta, K., Norris, S.A. and Cameron, N. (2014). Multi-level modelling of longitudinal child growth data from the Birth-to-Twenty Cohort: a comparison of growth models. *Annals of Human Biology*, 41(2):168–179.

Clarke, S. and Hall, P. (2009). Robustness of multiple testing procedures against dependence. *The Annals of Statistics*, 37(1):332–358.

Cole, T.J., Donaldson, M.D. and Ben-Shlomo, Y. (2010). SITAR – a useful instrument for growth curve analysis. *International Journal of Epidemiology*, 39(6):1558–1566.

Crookston, B.T., Schott, W., Cueto, S., Dearden, K.A., Engle, P., Georgiadis, A., Lundeen, E.A., Penny, M.E., Stein, A.D. and Behrman, J.R. (2013). Postinfancy growth, schooling, and cognitive achievement: Young lives. *The American Journal of Clinical Nutrition*, 98(6):1555–1563.

De Cao, E. (2015). The height production function from birth to age two. *Journal of Human Capital*, 9(3):329–363.

Durlauf, S.N., Johnson, P.A. and Temple, J.R. (2005). Growth Econometrics. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, volume 1, chapter 8, pp. 555–677. Elsevier.

Engle, P.L., Castle, S. and Menon, P. (1996). Child development: vulnerability and resilience. *Social Science & Medicine*, 43(5):621–635.

Fink, G. and Rockers, P.C. (2014). Childhood growth, schooling, and cognitive development: further evidence from the Young Lives study. *The American Journal of Clinical Nutrition*, 100(1):182–188.

Gandhi, M., Ashorn, P., Maleta, K., Teivaanmäki, T., Duan, X. and Cheung, Y.B. (2011). Height gain during early childhood is an important predictor of schooling and mathematics ability outcomes. *Acta Paediatrica*, 100(8):1113–1118.

Georgiadis, A., Benny, L., Crookston, B.T., Hermida, P., Mani, S., Woldehanna, T., Stein, A.D., Behrman, J.R. et al. (2016). Growth trajectories from conception through middle childhood and cognitive achievement at age 8 years: evidence from four low-and middle-income countries. *Socal Science & Medicine: Population Health*, 2:43–54.

Georgiadis, A., Benny, L., Galab, S., Reddy, P., Woldehanna, T. et al. (2017). Growth recovery and faltering though early adolescence in low-and middle-income countries: Determinants and implications for cognitive development. *Social Science & Medicine*, 179:81–90.

Guimarães, P. and Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. *Stata Journal*, 10(4):628–649.

Gurka, M.J., Edwards, L.J. and Muller, K.E. (2011). Avoiding bias in mixed model inference for fixed effects. *Statistics in Medicine*, 30(22):2696–2707.

Handa, S. and Peterman, A. (2016). Is there catch-up growth? Evidence from three continents. *Oxford Bulletin of Economics and Statistics*, 78(4):470–500.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica*, 46(6):1251–1272.

Heckman, J.J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902.

—— (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33):13250–13255.

Hirvonen, K. (2014). Measuring catch-up growth in malnourished populations. *Annals of Human Biology*, 41(1):67–75.

Hoddinott, J., Behrman, J.R., Maluccio, J.A., Melgar, P., Quisumbing, A.R., Ramirez-Zea, M., Stein, A.D., Yount, K.M. and Martorell, R. (2013). Adult consequences of growth failure in early childhood. *The American Journal of Clinical Nutrition*, 98(5):1170.

Hoddinott, J. and Kinsey, B. (2001). Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, 63(4):409–436.

Jacqmin-Gadda, H., Sibillot, S., Proust, C., Molina, J.M. and Thiébaut, R. (2007). Robustness of the linear mixed model to misspecified error distribution. *Computational Statistics & Data Analysis*, 51(10):5142–5154.

Jayachandran, S. and Pande, R. (2017). Why are Indian children so short? The role of birth order and son preference. *American Economic Review*, 107(9):2600–2629. doi:10.1257/aer.20151282.

Jerrim, J. and Vignoles, A. (2013). Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(4):887–906.

Johnson, W. (2015). Analytical strategies in human growth research. *American Journal of Human Biology*, 27(1):69–83.

Koedel, C., Mihaly, K. and Rockoff, J.E. (2015). Value-added modeling: A review. *Economics of Education Review*, 47:180–195.

Kuklina, E.V., Ramakrishnan, U., Stein, A.D., Barnhart, H.H. and Martorell, R. (2006). Early childhood growth and development in rural Guatemala. *Early Human Development*, 82(7):425–433.

Le, T.D. and Behrman, J.R. (2017). Heterogeneity in predictive power of early childhood nutritional indicators for mid-childhood outcomes: Evidence from Vietnam. *Economics & Human Biology*, 26:86–95.

Leroy, J.L., Ruel, M., Habicht, J.P. and Frongillo, E.A. (2014). Linear growth deficit continues to accumulate beyond the first 1000 days in low-and middle-income countries: global evidence from 51 national surveys. *The Journal of Nutrition*, 144(9):1460–1466.

—— (2015). Using height-for-age differences (HAD) instead of height-for-age z-scores (HAZ) for the meaningful measurement of population-level catch-up in linear growth in children less than 5 years of age. *BMC Pediatrics*, 15(1):145.

Lewis, J.B. and Linzer, D.A. (2005). Estimating regression models in which the dependent variable is based on estimates. *Political Analysis*, 13(4):345–364.

Lohman, D.F. and Korb, K.A. (2006). Gifted today but not tomorrow? Longitudinal changes in ability and achievement during elementary school. *Journal for the Education of the Gifted*, 29(4):451–484.

Lundeen, E.A., Behrman, J.R., Crookston, B.T., Dearden, K.A., Engle, P., Georgiadis, A., Penny, M.E. and Stein, A.D. (2014a). Growth faltering and recovery in children aged 1–8 years in four low-and middle-income countries: Young Lives. *Public Health Nutrition*, 17(9):2131–2137.

Lundeen, E.A., Stein, A.D., Adair, L.S., Behrman, J.R., Bhargava, S.K., Dearden, K.A., Gigante, D., Norris, S.A., Richter, L.M., Fall, C.H. et al. (2014b). Height-for-age z scores increase despite increasing height deficits among children in 5 developing countries. *The American journal of Clinical Nutrition*, 100(3):821–825.

Mani, S. (2012). Is there complete, partial, or no recovery from childhood malnutrition? empirical evidence from Indonesia. *Oxford Bulletin of Economics and Statistics*, 74(5):691–715.

McCrory, C., O'Leary, N., Fraga, S., Ribeiro, A., Barros, H., Kartiosuo, N., Raitakari, O., Kivimäki, M., Vineis, P., Layte, R. et al. (2017). Socioeconomic differences in children's growth trajectories from infancy to early adulthood: evidence from four European countries. *Journal of Epidemiology and Community Health*, 71(10):981.

de Onis, M., Blössner, M. and Borghi, E. (2012). Prevalence and trends of stunting among pre-school children, 1990–2020. *Public Health Nutrition*, 15(1):142–148.

Outes, I. and Porter, C. (2013). Catching up from early nutritional deficits? Evidence from rural Ethiopia. *Economics & Human Biology*, 11(2):148–163.

Polachek, S.W. and Kim, M.K. (1994). Panel estimates of the gender earnings gap: individual-specific intercept and individual-specific slope models. *Journal of Econometrics*, 61(1):23–42.

Ribeiro, E.P. (2007). The dynamics of firm size distribution. *Brazilian Review of Econometrics*, 27(2):199–222.

Rieger, M. and Trommlerová, S.K. (2016). Age-specific correlates of child growth. *Demography*, 53(1):241–267.

Roback, P.J. and Askins, R.A. (2005). Judicious use of multiple hypothesis tests. *Conservation Biology*, 19(1):261–267.

Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N. and Grantham-McGregor, S. (2015). The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in Bogota. *Journal of Human Resources*, 50(2):464–483.

Schott, W., Aurino, E., Penny, M.E. and Behrman, J.R. (2017). Adolescent mothers' anthropometrics and grandmothers' schooling predict infant anthropometrics in Ethiopia, India, Peru, and Vietnam. *Annals of the New York Academy of Sciences*, forthcoming.

Schott, W.B., Crookston, B.T., Lundeen, E.A., Stein, A.D., Behrman, J.R. et al. (2013). Periods of child growth up to age 8 years in Ethiopia, India, Peru and Vietnam: key distal household and community factors. *Social Science & Medicine*, 97(2013):278–287.

Shrimpton, R., Victora, C.G., de Onis, M., Lima, R.C., Blössner, M. and Clugston, G. (2001). Worldwide timing of growth faltering: implications for nutritional interventions. *Pediatrics*, 107(5):e75–e75.

Spencer, P.R., Sanders, K.A. and Judge, D.S. (2017). Growth curves and the international standard: How children's growth reflects challenging conditions in rural Timor-Leste. *American Journal of Physical Anthropology*, forthcoming.

Stoel, R.D. and Van Den Wittenboer, G. (2003). Time dependence of growth parameters in latent growth curve models with time invariant covariates. *Methods of Psychological Research*, 8(1):21–41.

Tanner, J.M. (1962). The evaluation of growth and maturity in children. In F. Gross (Ed.), *Protein Metabolism: Influence of Growth Hormone, Anabolic Steroids, and Nutrition in Health and Disease. An International Symposium Leyden, 25th–29th June, 1962*, pp. 361–382. Berlin, Heidelberg: Springer. ISBN 978-3-642-53147-7.

Teivaanmäki, T., Cheung, Y.B., Kortekangas, E., Maleta, K. and Ashorn, P. (2015). Transition between stunted and nonstunted status: both occur from birth to 15 years of age in Malawi children. *Acta Paediatrica*, 104(12):1278–1285.

Tu, Y.K. and Gilthorpe, M.S. (2007). Revisiting the relation between change and initial value: a review and evaluation. *Statistics in Medicine*, 26(2):443–457.

—— (2011). *Statistical thinking in epidemiology*. Florida: CRC Press, Taylor & Francis.

Ulijaszek, S.J. and Kerr, D.A. (1999). Anthropometric measurement error and the assessment of nutritional status. *British Journal of Nutrition*, 82(3):165–177.

Victora, C., Onis, M.d., Hallal, P., Blössner, M., Shrimpton, R. et al. (2010). Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics*, 125(3):e473–e480.

Vidmar, S., Carlin, J., Hesketh, K. and Cole, T. (2004). Standardizing anthropometric measures in children and adolescents with new functions for egen. *Stata Journal*, 4(1):50–55.

Vidmar, S.I., Cole, T.J. and Pan, H. (2013). Standardizing anthropometric measures in children and adolescents with functions for egen: Update. *Stata Journal*, 13(2):366–378.

Wilkinson, B. (1951). A statistical consideration in psychological research. *Psychological Bulletin*, 48(2):156.

Wit, J.M. and Boersma, B. (2002). Catch-up growth: definition, mechanisms, and models. *Journal of Pediatric Endocrinology and Metabolism*, 15:1229–1242.

Zhang, R., Undurraga, E.A., Zeng, W., Reyes-García, V., Tanner, S., Leonard, W.R., Behrman, J.R. and Godoy, R.A. (2016). Catch-up growth and growth deficits: Nine-year annual panel child growth for native Amazonians in Bolivia. *Annals of Human Biology*, 43(4):304–315.

A Appendix: Supplementary material

A.1 Tables

Table A1: Summary of properties of alternative latent growth regression estimators

Estimator	Key assumptions	Strengths	Weaknesses
POLS	$\sigma_\alpha = \sigma_\beta = \rho_{\alpha\beta} = 0$	Simple	Restrictive assumptions do not permit direct tests of within-group catch-up
CRE	$\rho_{x\beta} = \rho_{x\alpha} = 0$ $\alpha_i \sim \mathcal{N}(0, \sigma^2_\alpha)$ $\beta_i \sim \mathcal{N}(0, \sigma^2_\beta)$	Permits direct tests of all catch-up hypotheses	Random effects assumed to follow normal distribution; practical restrictions on covariance structure; convergence can fail / be slow
FEIS	–	Minimal assumptions; controls for all time invariant factors on size and growth velocity; rapid convergence	Measurement error in fixed effects; no direct tests of individual time invariant factors

Note: 'Assumptions' refer to the practical restrictions applied to equations (6a)–(6c) in standard applications of the different estimators.

Table A2: Estimates of catch-up growth in HAZ scores (unconditional model)

					Between			Within		Both
		α_0	β_0	$\rho_{\alpha\beta}$	$H0_{\alpha_0}$	$H0_{\beta_0}$	?	$H0_{\rho_{\alpha\beta}}$	?	?
Ethiopia	POLS	-1.383	0.005	.	0.00	0.96	N	.	.	.
		(0.023)	(0.003)							
	CRE	-1.389	0.007	-0.473	0.00	0.01	Y	0.00	Y	Y
		(0.023)	(0.003)	(0.023)						
	FEIS	-1.381	0.003	-0.297	0.00	0.20	N	0.00	Y	N
		(0.023)	(0.003)	(0.023)						
India	POLS	-1.438	-0.011	.	0.00	0.99	N	.	.	.
		(0.021)	(0.003)							
	CRE	-1.447	-0.009	-0.275	0.00	0.99	N	0.00	Y	N
		(0.022)	(0.003)	(0.023)						
	FEIS	-1.439	-0.007	-0.191	0.00	0.99	N	0.00	Y	N
		(0.021)	(0.003)	(0.023)						
Peru	POLS	-1.275	0.041	.	0.00	0.99	N	.	.	.
		(0.022)	(0.002)							
	CRE	-1.284	0.041	-0.073	0.00	0.00	Y	0.00	Y	Y
		(0.023)	(0.002)	(0.023)						
	FEIS	-1.273	0.039	-0.029	0.00	0.00	Y	0.10	N	N
		(0.022)	(0.002)	(0.023)						
Vietnam	POLS	-1.126	0.005	.	0.00	0.99	N	.	.	.
		(0.022)	(0.002)							
	CRE	-1.129	0.006	0.056	0.00	0.00	Y	0.99	N	N
		(0.022)	(0.002)	(0.023)						
	FEIS	-1.126	0.016	0.036	0.00	0.00	Y	0.94	N	N
		(0.022)	(0.002)	(0.023)						

Note: see Table 4; no additional covariates included in the models; probability estimates larger than 0.99 have been rounded down to the latter value.
Source: own estimates.

Table A3: Goodness-of-fit metrics, HAZ scores

	Unconditional				Conditional			
	Eth.	Ind.	Peru	Viet.	Eth.	Ind.	Peru	Viet.
POLS estimates:								
Time	0.04	0.01	0.01	0.01	0.33	0.19	0.16	0.19
RMSE	1.25	1.13	1.12	1.11	1.21	1.06	0.94	0.98
Adj. R^2	0.00	0.00	0.02	0.00	0.07	0.11	0.28	0.21
AIC	2.4e+04	2.3e+04	2.3e+04	2.3e+04	2.3e+04	2.2e+04	1.7e+04	2.1e+04
CRE estimates:								
Time	6.88	6.64	6.86	7.22	58.41	62.07	53.87	116.30
RMSE	0.64	0.49	0.44	0.40	0.52	0.39	0.33	0.40
AIC	2.2e+04	2.0e+04	1.8e+04	1.8e+04	2.1e+04	1.9e+04	1.4e+04	1.7e+04
FEIS estimates :								
Time	0.41	0.46	0.58	0.52	0.46	0.55	0.68	0.58
RMSE	0.54	0.41	0.37	0.33	0.54	0.40	0.30	0.33
Adj. R^2	0.62	0.74	0.78	0.82	0.62	0.74	0.81	0.82
AIC	1.2e+04	7.9e+03	6.4e+03	4.8e+03	1.2e+04	7.8e+03	2.9e+03	4.7e+03
Hausman	.	.	.	.	0.00	0.00	0.00	0.00

Note: Time is the execution speed of the regression command, in seconds; Hausman gives the probability associated with a Hausman test where the null hypothesis is that the CRE estimator is consistent (and efficient) versus the alternative hypothesis which is that the FEIS estimator is consistent.
Source: own estimates.

Table A4: Size and velocity regressions, Ethiopia

	Height (HAZ)					BMI (BAZ)				
	Size	Size	Velocity	Velocity	Velocity	Size	Size	Velocity	Velocity	Velocity
Mean	-5.781***	-5.774***	-0.129***	-0.129***	-0.129***	-8.988***	-8.998***	-1.328***	-1.326***	-1.327***
	(0.074)	(0.074)	(0.009)	(0.010)	(0.008)	(0.107)	(0.121)	(0.021)	(0.024)	(0.022)
Female	0.375**	0.412***	-0.027	-0.020	-0.019	-1.429***	-1.475***	-0.111***	-0.126***	-0.150***
	(0.155)	(0.148)	(0.019)	(0.018)	(0.019)	(0.212)	(0.247)	(0.040)	(0.037)	(0.040)
Urban	-0.265	1.096	0.014	0.111	0.115	-0.102	1.330	0.105	0.324	0.345
	(0.258)	(0.832)	(0.031)	(0.146)	(0.136)	(0.368)	(1.681)	(0.067)	(0.335)	(0.294)
Not first born	-0.032	-0.122	-0.018	-0.002	-0.002	-0.535	-0.273	0.187***	0.177***	0.173***
	(0.217)	(0.190)	(0.028)	(0.026)	(0.032)	(0.354)	(0.390)	(0.063)	(0.055)	(0.056)
Majority ELG	0.358*	0.257	-0.038	0.027	0.028	-1.236***	-0.179	-0.260***	-0.084	-0.086
	(0.183)	(0.310)	(0.023)	(0.035)	(0.033)	(0.301)	(0.433)	(0.054)	(0.081)	(0.075)
Wealth index	1.143***	1.147***	-0.005	-0.024	-0.020	-0.089	-0.569**	-0.116***	-0.083**	-0.092**
	(0.130)	(0.130)	(0.018)	(0.019)	(0.017)	(0.196)	(0.244)	(0.036)	(0.038)	(0.039)
Mother's edu.	0.228*	0.186	-0.038**	-0.029*	-0.028*	0.168	-0.196	0.000	0.025	0.022
	(0.117)	(0.121)	(0.016)	(0.016)	(0.017)	(0.204)	(0.231)	(0.038)	(0.037)	(0.037)
Mother's age	0.004	0.006	-0.004**	-0.004**	-0.004**	0.020	0.017	0.004	0.002	0.002
	(0.014)	(0.015)	(0.002)	(0.002)	(0.002)	(0.027)	(0.023)	(0.004)	(0.004)	(0.004)
Sibs 0-5	-0.495**	-0.492**	-0.104***	-0.057**	-0.059**	-0.647**	-0.399	-0.027	-0.198***	-0.204***
	(0.212)	(0.195)	(0.025)	(0.026)	(0.025)	(0.320)	(0.326)	(0.056)	(0.050)	(0.050)
Sibs 6-12	-0.063	-0.054	-0.062**	-0.051*	-0.052*	-0.700**	-0.796**	0.028	0.025	0.012
	(0.178)	(0.196)	(0.029)	(0.027)	(0.027)	(0.327)	(0.332)	(0.059)	(0.048)	(0.053)
Child FE (std.)	-	-	-	-	-0.013	-	-	-	-	-0.086***
					(0.011)					(0.029)
Obs.	1,845	1,845	1,845	1,845	1,845	1,846	1,846	1,846	1,846	1,846
RMSE	3.10	3.03	0.39	0.38	0.38	5.23	5.01	0.91	0.82	0.81
Adj. R^2	0.13	0.17	0.02	0.09	0.09	0.05	0.13	0.04	0.22	0.23
WI$\times$2 + Mean	-3.50	-3.48	-0.14	-0.18	-0.17	-9.17	-10.14	-1.56	-1.49	-1.51
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.

Source: own estimates.

Table A5: Size and velocity regressions, India

	Height (HAZ)					BMI (BAZ)				
	Size	Size	Velocity	Velocity	Velocity	Size	Size	Velocity	Velocity	Velocity
Mean	-6.061***	-6.061***	-0.222***	-0.222***	-0.222***	-9.682***	-9.681***	-0.224***	-0.223***	-0.222***
	(0.065)	(0.066)	(0.009)	(0.009)	(0.008)	(0.178)	(0.139)	(0.036)	(0.039)	(0.033)
Female	0.272**	0.195	0.010	0.011	0.008	-0.087	-0.138	0.061	0.044	0.067
	(0.132)	(0.131)	(0.017)	(0.015)	(0.017)	(0.272)	(0.257)	(0.069)	(0.073)	(0.059)
Urban	-0.163	-0.112	0.065***	-0.024	-0.023	0.730	-0.843	0.123	-0.340	-0.200
	(0.203)	(0.674)	(0.024)	(0.087)	(0.094)	(0.500)	(1.569)	(0.091)	(0.315)	(0.298)
Not first born	-0.476***	-0.390**	-0.048**	-0.047**	-0.042*	-0.249	-0.326	-0.048	-0.071	-0.017
	(0.165)	(0.164)	(0.022)	(0.022)	(0.022)	(0.333)	(0.385)	(0.069)	(0.090)	(0.059)
Majority ELG	-0.255**	-0.133	0.000	-0.015	-0.014	-1.010***	-0.540	-0.284***	-0.169***	-0.079
	(0.130)	(0.140)	(0.017)	(0.017)	(0.014)	(0.311)	(0.329)	(0.072)	(0.057)	(0.058)
Wealth index	0.662***	0.642***	0.047***	0.060***	0.052***	1.458***	1.312***	0.088	0.072	-0.146
	(0.101)	(0.108)	(0.011)	(0.013)	(0.013)	(0.250)	(0.245)	(0.067)	(0.065)	(0.126)
Mother's edu.	0.223**	0.170*	-0.002	-0.008	-0.010	0.487	0.487	0.144	0.139	0.059
	(0.107)	(0.099)	(0.011)	(0.012)	(0.013)	(0.366)	(0.319)	(0.101)	(0.099)	(0.038)
Mother's age	0.046***	0.050***	-0.006***	-0.006**	-0.006***	-0.006	-0.000	-0.002	-0.002	-0.002
	(0.015)	(0.018)	(0.002)	(0.002)	(0.002)	(0.036)	(0.033)	(0.008)	(0.008)	(0.007)
Sibs 0-5	-0.038	0.029	0.048*	0.051**	0.051**	-0.068	-0.116	-0.102	-0.148	-0.128
	(0.223)	(0.217)	(0.027)	(0.025)	(0.025)	(0.587)	(0.538)	(0.108)	(0.146)	(0.104)
Sibs 6-12	-0.015	0.090	0.019	0.003	0.002	-0.491	-0.189	-0.132**	-0.092	-0.061
	(0.195)	(0.196)	(0.023)	(0.023)	(0.019)	(0.425)	(0.380)	(0.063)	(0.065)	(0.070)
Child FE (std.)	-	-	-	-	0.040***	-	-	-	-	1.260**
					(0.009)					(0.511)
Obs.	1,913	1,913	1,913	1,913	1,913	1,913	1,913	1,913	1,913	1,913
RMSE	2.90	2.86	0.35	0.33	0.33	7.30	7.28	1.79	1.79	1.32
Adj. R^2	0.08	0.11	0.05	0.13	0.14	0.07	0.08	0.02	0.03	0.47
WI$\times$2 + Mean	-4.74	-4.78	-0.13	-0.10	-0.12	-6.77	-7.06	-0.05	-0.08	-0.51
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.

Source: own estimates.

Table A6: Size and velocity regressions, Peru

	Height (HAZ)					BMI (BAZ)				
	Size	Size	Velocity	Velocity	Velocity	Size	Size	Velocity	Velocity	Velocity
Mean	-5.354***	-5.351***	0.098***	0.097***	0.097***	8.332***	8.336***	0.343***	0.343***	0.343***
	(0.079)	(0.071)	(0.009)	(0.009)	(0.010)	(0.181)	(0.209)	(0.036)	(0.037)	(0.034)
Female	-0.094	-0.125	0.038**	0.031	0.031	-1.335***	-1.333***	-0.092	-0.089	-0.038
	(0.125)	(0.145)	(0.019)	(0.019)	(0.021)	(0.391)	(0.376)	(0.070)	(0.086)	(0.072)
Urban	0.367*	-0.211	-0.022	0.011	0.011	-0.982*	-1.421**	0.164	0.050	0.106
	(0.212)	(0.257)	(0.030)	(0.038)	(0.032)	(0.558)	(0.657)	(0.104)	(0.130)	(0.133)
Not first born	0.045	-0.063	-0.033	-0.035	-0.035	-0.437	-0.715	-0.003	-0.021	0.007
	(0.155)	(0.189)	(0.028)	(0.025)	(0.025)	(0.555)	(0.516)	(0.100)	(0.087)	(0.093)
Majority ELG	0.355*	0.069	-0.043	-0.036	-0.036	-0.546	-0.099	0.117	0.153	0.157
	(0.187)	(0.167)	(0.027)	(0.032)	(0.027)	(0.572)	(0.532)	(0.105)	(0.129)	(0.121)
Wealth index	0.913***	0.824***	0.022*	0.048***	0.050***	2.107***	1.619***	0.362***	0.248***	0.185***
	(0.084)	(0.100)	(0.012)	(0.018)	(0.015)	(0.248)	(0.279)	(0.045)	(0.061)	(0.056)
Mother's edu.	0.530***	0.471***	-0.010	-0.008	-0.007	0.435	0.454*	-0.012	0.001	-0.017
	(0.091)	(0.084)	(0.013)	(0.014)	(0.014)	(0.302)	(0.269)	(0.054)	(0.061)	(0.059)
Mother's age	0.019*	0.024**	-0.004**	-0.004**	-0.004**	0.040	0.046	-0.013**	-0.011*	-0.013*
	(0.011)	(0.011)	(0.002)	(0.002)	(0.002)	(0.037)	(0.032)	(0.006)	(0.006)	(0.006)
Sibs 0-5	0.076	0.087	-0.076***	-0.076***	-0.076***	0.003	-0.140	-0.270***	-0.289***	-0.283***
	(0.167)	(0.172)	(0.024)	(0.027)	(0.023)	(0.508)	(0.450)	(0.102)	(0.105)	(0.093)
Sibs 6-12	-0.669***	-0.496***	0.023	-0.002	-0.003	-0.523	0.063	0.107	0.154	0.151
	(0.192)	(0.184)	(0.035)	(0.027)	(0.031)	(0.512)	(0.457)	(0.105)	(0.104)	(0.116)
Child FE (std.)	-	-	-	-	-0.007	-	-	-	-	0.354***
					(0.015)					(0.057)
Obs.	1,916	1,916	1,916	1,916	1,916	1,921	1,921	1,921	1,921	1,921
RMSE	2.85	2.78	0.42	0.41	0.41	8.67	8.28	1.63	1.62	1.59
Adj. R^2	0.31	0.35	0.02	0.05	0.05	0.10	0.17	0.11	0.13	0.16
WI×2 + Mean	-3.53	-3.70	0.14	0.19	0.20	12.55	11.57	1.07	0.84	0.71
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.
Source: own estimates.

Table A7: Size and velocity regressions, Vietnam

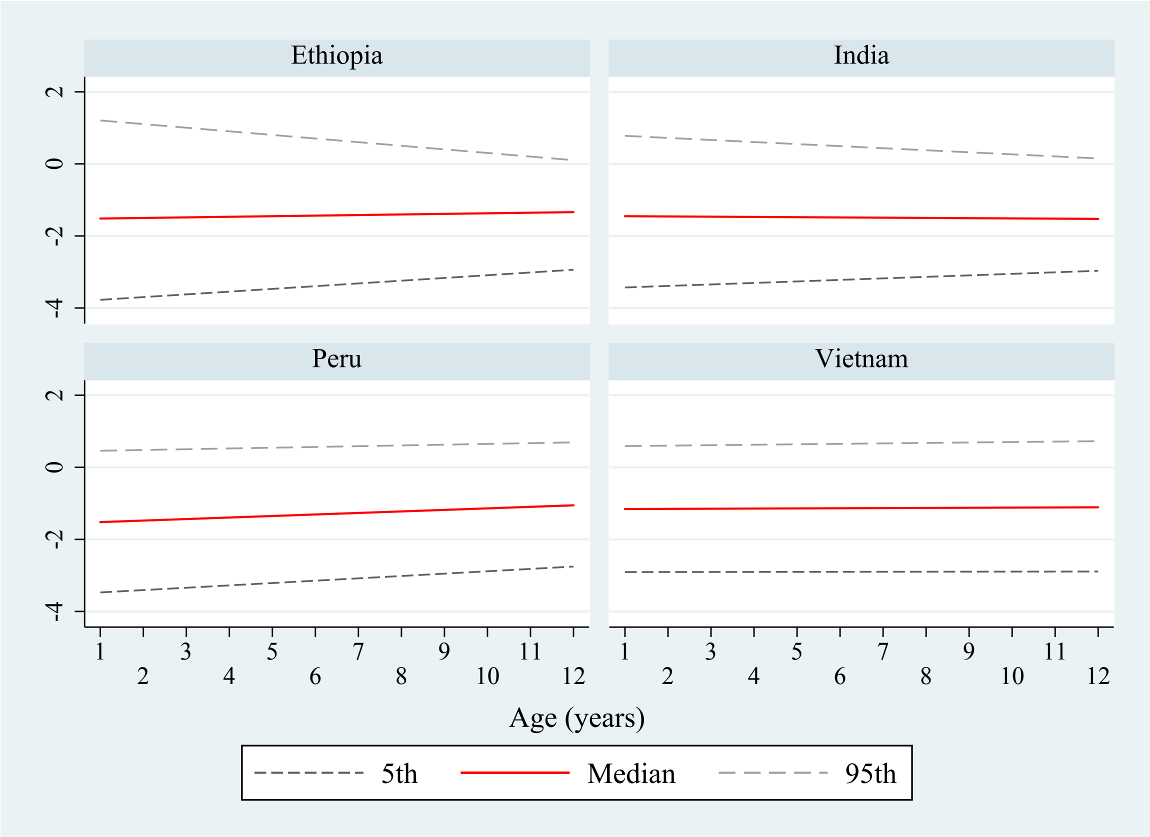
	Height (HAZ)					BMI (BAZ)				
	Size	Size	Velocity	Velocity	Velocity	Size	Size	Velocity	Velocity	Velocity
Mean	-4.709***	-4.709***	-0.012*	-0.012*	-0.012*	-3.433***	-3.429***	-0.181***	-0.181***	-0.182***
	(0.060)	(0.071)	(0.007)	(0.006)	(0.007)	(0.157)	(0.183)	(0.026)	(0.026)	(0.023)
Female	0.187	0.139	0.005	0.006	0.001	-1.764***	-1.720***	-0.159***	-0.162***	-0.016
	(0.123)	(0.135)	(0.014)	(0.013)	(0.013)	(0.348)	(0.343)	(0.049)	(0.048)	(0.039)
Urban	1.313***	0.843	-0.001	0.013	-0.016	4.646***	6.147***	0.177**	0.195	-0.327
	(0.213)	(0.653)	(0.023)	(0.087)	(0.090)	(0.602)	(1.826)	(0.076)	(0.279)	(0.235)
Not first born	-0.451**	-0.383**	-0.019	-0.017	-0.004	-1.131**	-1.125**	-0.154**	-0.138**	-0.043
	(0.182)	(0.163)	(0.019)	(0.020)	(0.018)	(0.478)	(0.477)	(0.070)	(0.063)	(0.061)
Majority ELG	1.272***	0.294	0.005	-0.050	-0.060	-1.473***	2.536**	0.013	-0.285*	-0.500***
	(0.183)	(0.318)	(0.019)	(0.053)	(0.052)	(0.452)	(1.053)	(0.062)	(0.171)	(0.176)
Wealth index	0.648***	0.893***	0.054***	0.057***	0.027**	0.948***	1.550***	0.226***	0.291***	0.159***
	(0.107)	(0.108)	(0.012)	(0.014)	(0.012)	(0.266)	(0.311)	(0.040)	(0.045)	(0.036)
Mother's edu.	0.567***	0.580***	0.023**	0.014	-0.006	0.815***	0.808**	-0.010	-0.002	-0.071**
	(0.108)	(0.104)	(0.010)	(0.012)	(0.013)	(0.313)	(0.316)	(0.044)	(0.043)	(0.036)
Mother's age	-0.017	-0.024	-0.006***	-0.005***	-0.004**	0.040	-0.001	0.012**	0.009*	0.009**
	(0.014)	(0.015)	(0.001)	(0.001)	(0.002)	(0.037)	(0.034)	(0.005)	(0.005)	(0.004)
Sibs 0-5	-0.895***	-0.459	-0.026	-0.013	0.003	0.528	0.324	-0.136	-0.091	-0.119
	(0.284)	(0.293)	(0.031)	(0.034)	(0.030)	(0.687)	(0.729)	(0.100)	(0.088)	(0.097)
Sibs 6-12	-0.434	0.022	-0.055**	-0.046*	-0.046**	-1.097	0.212	-0.230***	-0.104	-0.122
	(0.267)	(0.289)	(0.027)	(0.026)	(0.023)	(0.684)	(0.638)	(0.087)	(0.090)	(0.087)
Child FE (std.)	-	-	-	-	0.115***	-	-	-	-	0.684***
					(0.010)					(0.028)
Obs.	1,908	1,908	1,908	1,908	1,908	1,909	1,909	1,909	1,909	1,909
RMSE	2.85	2.74	0.30	0.30	0.28	7.47	7.19	1.07	1.04	0.85
Adj. R^2	0.29	0.35	0.07	0.10	0.19	0.14	0.20	0.10	0.14	0.44
WI×2 + Mean	-3.41	-2.92	0.10	0.10	0.04	-1.54	-0.33	0.27	0.40	0.14
Community FE?	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

significance: * 10%, ** 5%, *** 1%.

Note: table reports selected coefficient estimates from auxiliary analyses, where the dependent variable is denoted in the column header and is derived from the FEIS estimates (Tables 5 and 6); individual children are the unit of observation and all covariates are taken as averages across rounds (demeaned); standard errors are based on 100 bootstrap iterations; community fixed effects are included as denoted in the final row.
Source: own estimates.

A.2 Figures

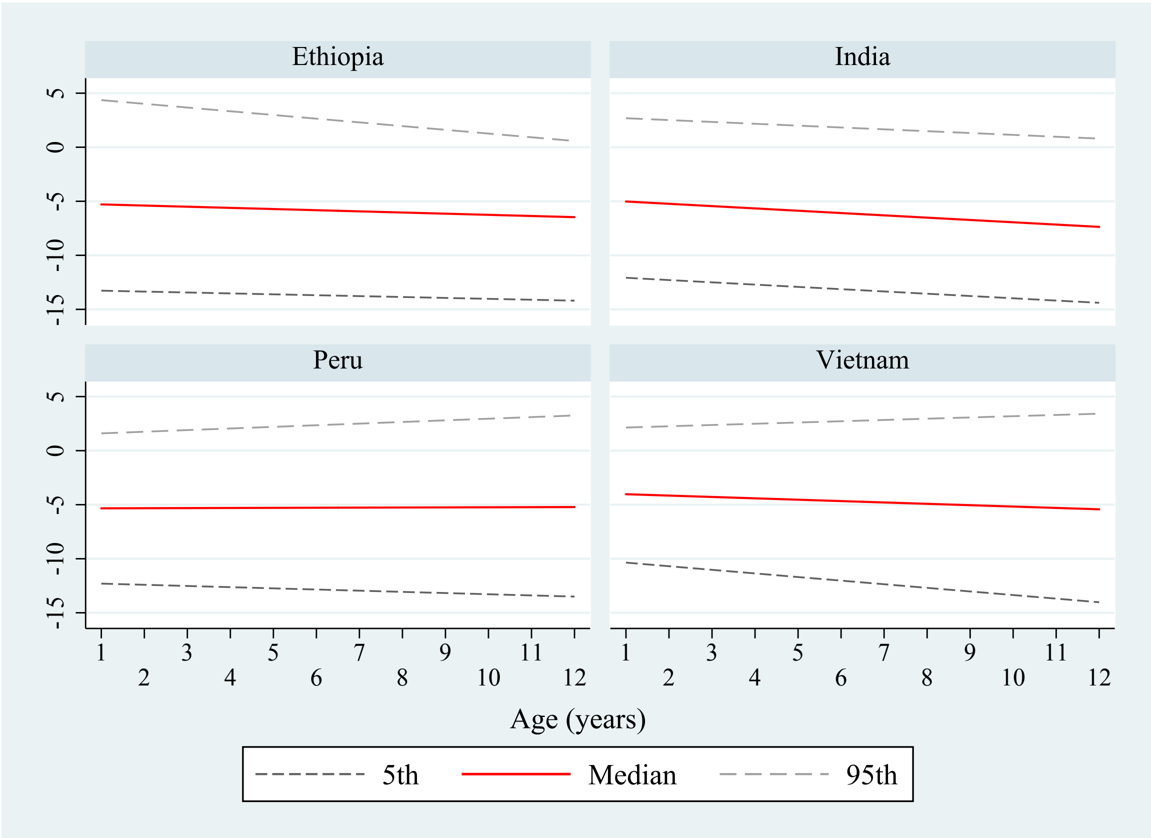
Figure A1: Trends in height-for-age *z*-scores (HAZ), by country



Notes: trends estimated via quantile regression.

Source: own estimates

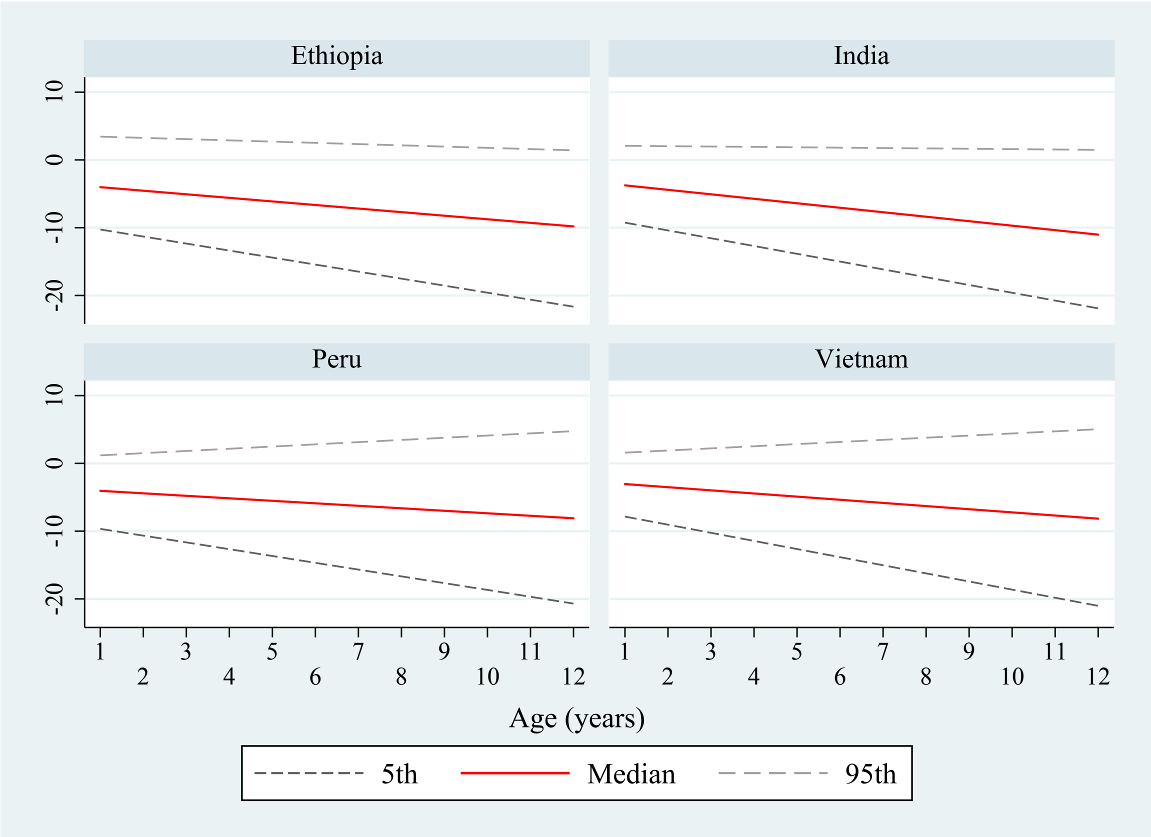
Figure A2: Trends in height-for-age ratios (HAR), by country



Notes: trends estimated via quantile regression.

Source: own estimates

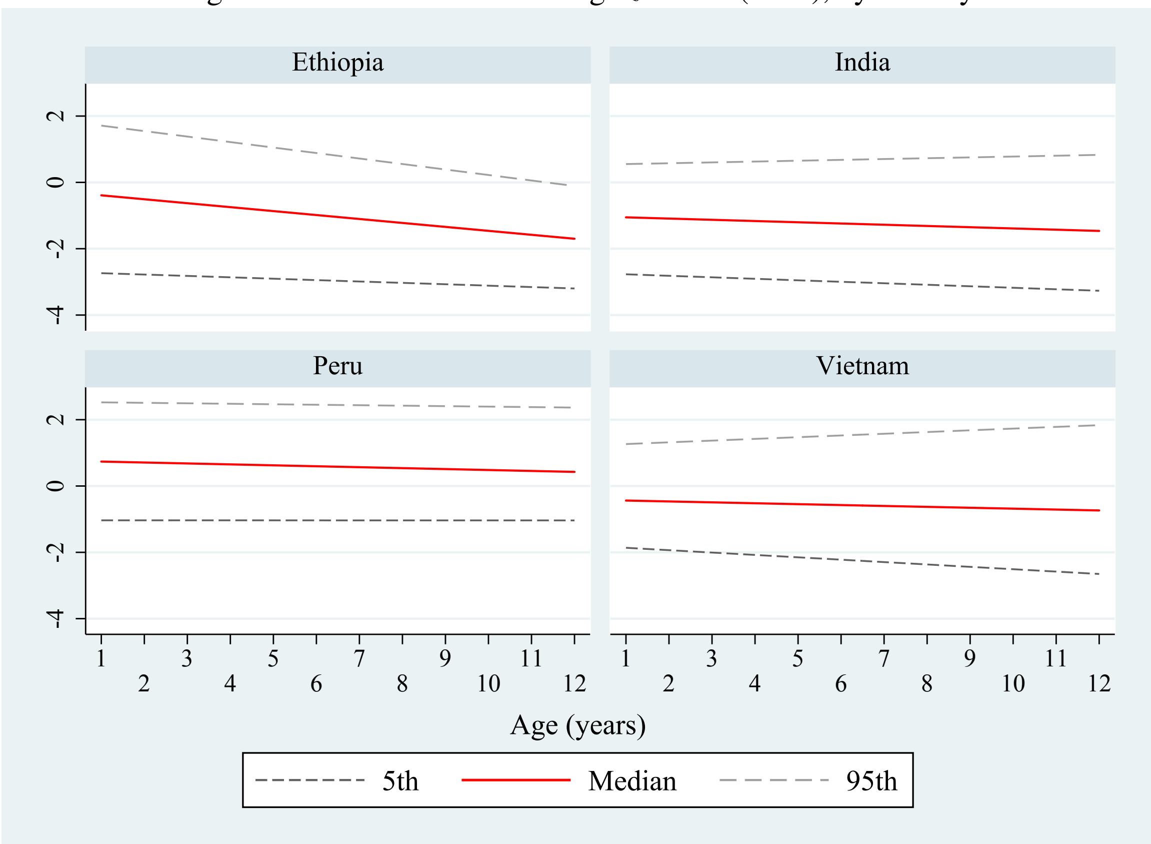
Figure A3: Trends in height-for-age differences (HAD), by country



Notes: trends estimated via quantile regression.

Source: own estimates

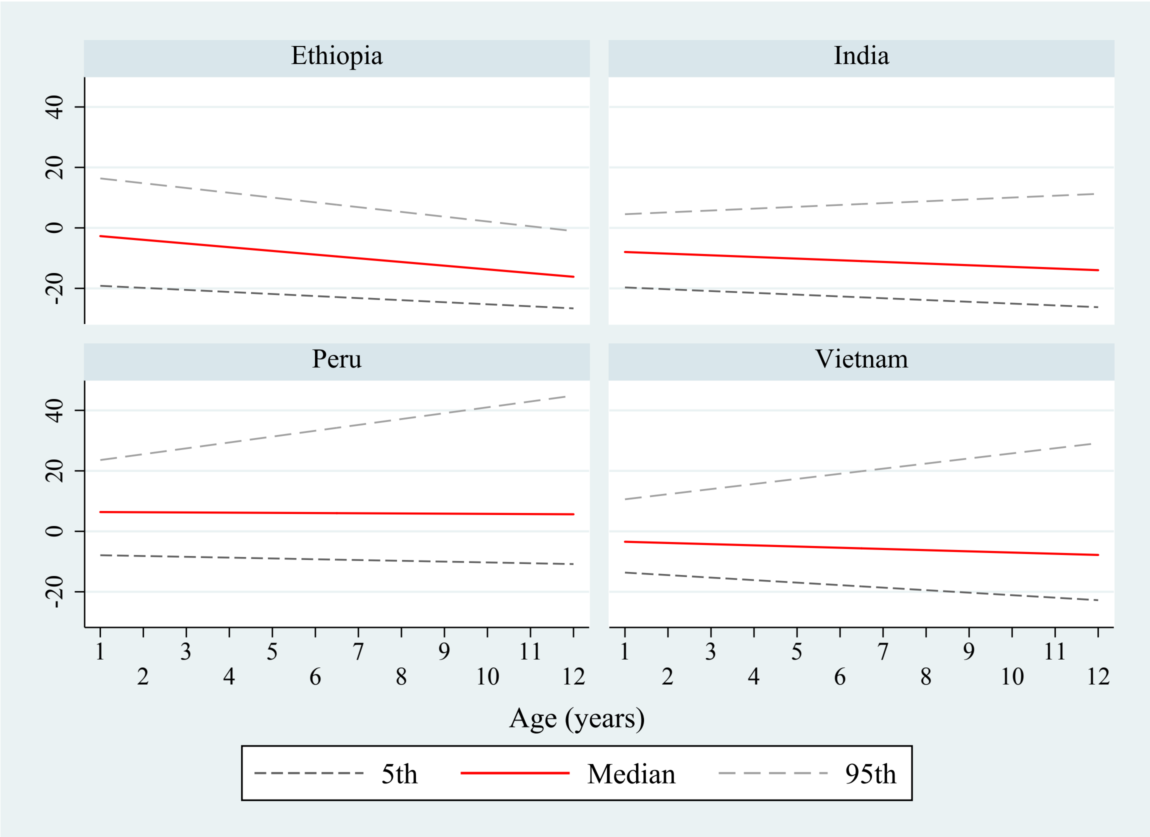
Figure A4: Trends in BMI-for-age *z*-scores (BAZ), by country



Notes: trends estimated via quantile regression.

Source: own estimates

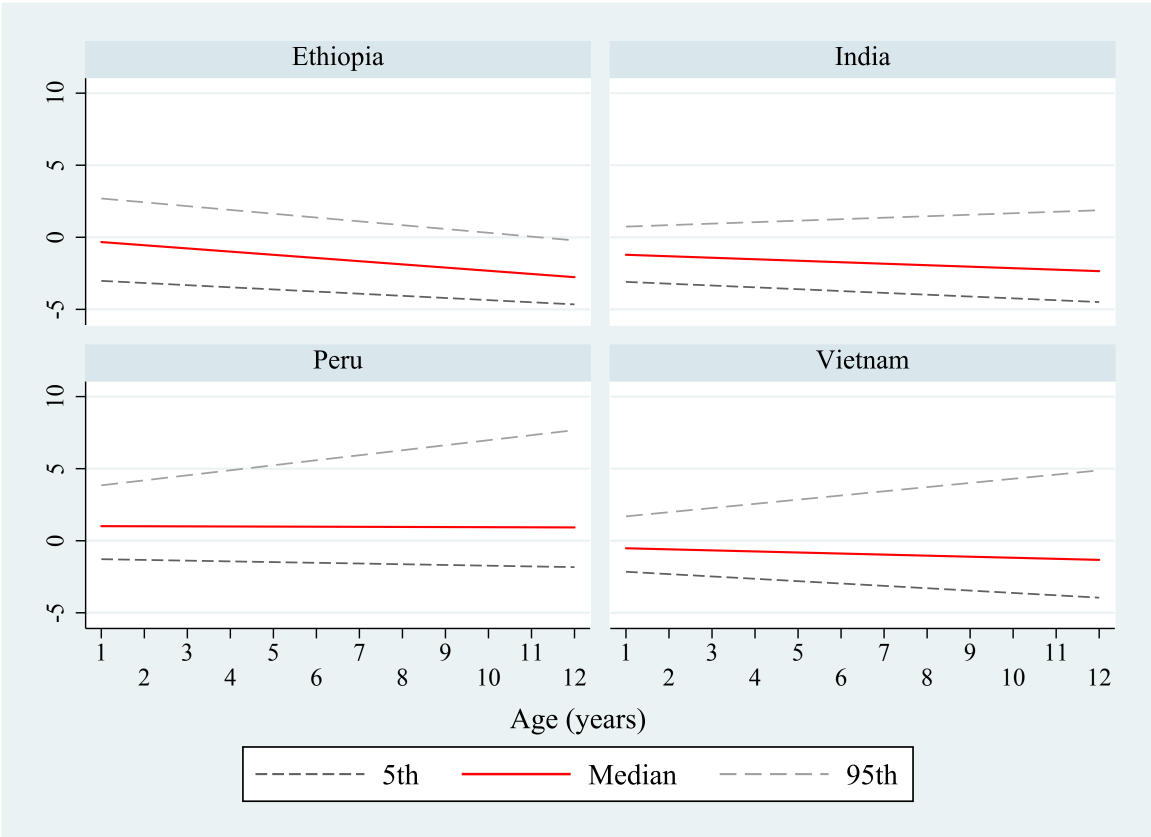
Figure A5: Trends in BMI-for-age ratios (BAR), by country



Notes: trends estimated via quantile regression.

Source: own estimates

Figure A6: Trends in BMI-for-age differences (BAD), by country



Notes: trends estimated via quantile regression.

Source: own estimates