ARTICLE IN PRESS

Social Science & Medicine xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Varied patterns of catch-up in child growth: Evidence from Young Lives

Paul Anand[a], Jere R. Behrman[b], Hai-Anh H. Dang[c], Sam Jones[d,*]

[a] Professor of Economics, Open University, UK
[b] William R. Kenan Jr. Professor of Economics, University of Pennsylvania, USA
[c] Survey Unit, Development Data Group, World Bank, USA
[d] Associate Professor, University of Copenhagen, Denmark

A R T I C L E I N F O

Keywords:
Ethiopia
India
Peru
Vietnam
Young lives
Catch-up growth
Height
Latent growth

JEL classification:
I12
J13

A B S T R A C T

The development of human capabilities for many disadvantaged children around the world depends on growth recovery ('catch-up growth'). Here we develop a novel framework that allows different types of catch-up growth to be classified and estimated. We distinguish between catch-up in the mean of a group toward that of a healthy reference population versus catch-up within the group. We show these different growth types can be tested in a unified setting using a latent growth framework. We apply the results to four developing countries, using longitudinal data on 7641 children collected over the period 2002–2013. The results show catch-up growth rates are generally modest but vary significantly between countries, and that local environmental factors are material to variation in child growth trajectories. The paper discusses the benefits of the new framework versus current methods, shows that the method is feasible, and suggests they call for intervention designs that are sensitive to community and country contexts.

1. Introduction

There is growing interest in the contribution of health to the development of human capabilities (Sen, 2002; Coast et al., 2008; Haisma et al., 2017; Anand and Roope, 2016). In particular, evidence shows that early childhood is of critical importance (Black et al., 2017; Britto et al., 2017; Richter et al., 2017; Heckman, 2007) and disadvantages experienced during this period can cast a long shadow, affecting outcomes over the full life course. Among a range of developmental risk factors, poor nutrition is a major concern in developing country contexts. Indeed, while the prevalence of certain infectious diseases and child mortality have shown steep declines across many countries, levels of stunting remain stubbornly high, especially in sub-Saharan Africa and parts of South Asia (de Onis et al., 2012). A key area of debate concerns the nature and scope of 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). Minimally, such insights should help inform resource allocation priorities.

Our point of departure is that extant definitions and associated statistical tests for catch-up growth among children are not entirely convincing. 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 2000 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 heated (e.g., Hirvonen, 2014). A first aim of this paper is to clarify how catch-up growth is defined. While existing studies use a range of different definitions, we distinguish between two primary 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). Critically, these two forms are not necessarily consistent. For instance, the overall mean can change without a corresponding change in

We acknowledge very helpful comments from two anonymous referees, 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.

* Corresponding author. Økonomisk Institut, Københavns Universitet, Øster Farimagsgade 5, bygning 26, 1353, København K, Denmark.
E-mail address: sam.jones@econ.ku.dk (S. Jones).

https://doi.org/10.1016/j.socscimed.2018.07.003
Received 25 February 2018; Received in revised form 18 June 2018; Accepted 2 July 2018

the relative position of children in the distribution. Also, if the anthropometric outcomes of *both* initially advantaged and initially disadvantaged children converge toward the group mean, then within-group catch-up may occur without between-group catch-up.

In addition to looking at various forms of catch-up, previous studies also express anthropometric outcomes in different ways. This refers to what kind of transformation or normalization is applied. For instance, catch-up growth has been studied using raw height, raw deviations of height from that of a healthy reference population, and height-for-age *z*-scores (also referenced against a healthy population). Since the choice of transformation can be material (Desmond and Casale, 2017), we contribute a simple four-way typology of ways in which catch-up growth can be studied, combining the two main forms of catch-up and two dominant metrics. Based on this typology, we note that existing studies in the literature have typically focussed on only one or other of the two forms; and few studies consider whether their results are sensitive to the way in which outcomes are expressed. As such, we suggest that previous studies have not provided comprehensive insights regarding recovery dynamics in child growth.

In light of these deficiencies, a second aim of this paper is to compare and contrast evidence regarding the presence and magnitude of these different types of catch-up growth. Our initial hypothesis is that different types of catch-up growth may be associated with disparate underlying dynamics or etiologies. Imagine a health program that identifies children who are severely stunted in a community and teaches families how this can be treated at home via changes to stunted children's diets. While this may lead to a reduction in the prevalence of severe stunting in the community, it is possible this comes at the expense of other children in the family and mean anthropometric outcomes in the community remain unchanged – e.g., if a fixed amount of food is redistributed within the family toward stunted children. By contrast, rising incomes or improved harvests across a community may generate both a reduction in stunting and improved anthropometric outcomes on average. This motivates the importance of distinguishing, both conceptually and empirically, between different *types* of catch-up growth. It also suggests the need to compare patterns across these types, including possible differences in underlying drivers.

The remainder of the paper is structured as follows: Section 2 draws on previous literature to set out a four-way classification of types of catch-up growth based on two main axes of differentiation: (i) within-group versus between-group catch-up; and (ii) raw versus relative outcome measures. Section 3 considers how these different types of catch-up may be identified empirically. We show that a latent growth model provides a unified testing framework, permitting separate estimates for within- and between-group catch-up in the same specification. Latent growth models have been used to quantify and explain (socio-economic) gradients in child growth outcomes (e.g., McCrory et al., 2017), but they have not been widely used to test for the presence of catch-up growth. Moreover, as we discuss, such models address a number of shortcomings associated with conventional empirical tests of catch-up growth.

In Section 4 we apply the approach, focussing on the four countries covered by the Young Lives surveys (Ethiopia, India, Peru and Vietnam). These are chosen not merely due to the availability of high quality and comparable data, but also because the surveys refer to generally disadvantaged communities in very different socio-economic and cultural contexts. In Supplementary Material we provide further technical details about the estimation of the latent growth model (Appendix A); and we extend the aggregate analysis to investigate systematic differences in the patterns and drivers of catch-up growth (Appendix B). Although this is somewhat limited in scope (and further research along these lines is planned), it allows us to identify some of the main time-invariant factors associated with variation in the strength and direction of catch-up. We find that the relevance of these different factors varies considerably among the four countries (also see Jones et al., 2018). Section 5 reflects on the findings and concludes.

2. Definitions

Existing literature on catch-up growth spans contributions across the medical, biological and social sciences. Despite the large number of studies, existing scholarship does not employ a unique or consistent definition of what actually constitutes catch-up growth. Looking across the different ways in which catch-up growth has been investigated, two main conceptual distinctions are found. The first concerns the nature or form of catch-up growth. Following Hirvonen (2014), it is useful to distinguish between what we call within- and between-group catch-up. Within-group catch-up focuses on whether initially disadvantaged children grow at a more rapid rate than other children in the given sample – i.e., they converge towards the mean of the group (also Wit and Boersma, 2002). Estimates of within-group catch-up thus are typically concerned with the stability of individual rankings in the outcome distribution, or how the *shape* of the outcome distribution evolves over time for a given sample. The second form of catch-up concerns whether the average child in a sampled group is converging toward the median outcome observed in a healthy reference population. Here the focus is on how the group mean evolves relative to the median of an external reference group. So, in contrast to the first form, this says nothing about any changes of individuals' positions within the sample.

As already noted, in addition to differences in forms of catch-up, anthropometric outcomes are expressed in various ways. Numerous studies investigate catch-up growth using deviations from the median of a well-nourished population of children of the same age. Some studies use raw deviations from the medians, measured, for example in centimetres (Desmond and Casale, 2017 refer to such metrics as ‘absolute’; however, to avoid confusion we prefer ‘raw’). For instance, Leroy et al. (2015) propose that catch-up should be evaluated using the raw differences in height-for-age from the median of a healthy reference population (denoted, HAD). Other studies use the same deviations from reference medians, but express them as relative to the standard deviation in the well-nourished population, yielding height-for-age *z*-scores (HAZ). Here the assumption is that a raw deviation of, say, 6 cm from the median of a healthy population means something different for a child of 6 months of age compared to a child of 6 years of age. Cameron et al. (2005) note that, since the dispersion of children's height tends to increase with age, variation in HAZ scores can be driven largely or entirely by changes in the denominator – i.e., even if the mean raw height difference between a sample and the median of a reference population (HAD) remains constant or increases, the absolute magnitude of the corresponding *z*-score may fall with age.

The two binary distinctions discussed above (within-between & raw-relative) combine to yield four separate types of catch-up growth. Although few studies make these definitional distinctions explicit, Table 1 applies the four-way classification to recent scholarship, where the different types of catch-up considered by each study are indicated by a tick (✓). Two points are immediately apparent from the table. First, no single or even predominant type of catch-up growth has been analysed – i.e., the literature is replete with different definitions of catch-up. Second, none of the studies consider all four definitions of catch-up together. In fact, with few exceptions (Fink and Rockers, 2014; Zhang et al., 2016), all of studies focus on either the between- or within-group types only. These insights are material. Logically (and as we go on to confirm later), there is no *necessary* relationship between any of the four types of catch-up growth – i.e., the presence of one type does not imply other types of catch-up will also obtain. Consequently, while specific definitions of catch-up respond to distinct research questions, it follows that existing studies provide limited empirical evidence regarding any similarities or differences between the various types of catch-up, either in individual contexts or across multiple contexts. As noted in the introduction, a broader empirical understanding of (variations in) patterns in catch-up growth may yield a deeper understanding of relevant sources of differences.

Before proceeding, one additional comment merits note. A number

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

Study	Context	Between		Within		Key findings
		Raw	Rel.	Raw	Rel.	
M2012	Indonesia	–	–	✓	–	Incomplete within catch-up in raw height
O2013	Ethiopia	–	–	–	✓	Incomplete within catch-up in HAZ
S2013	Ethiopia, India, Peru, Vietnam	–	–	–	✓	Some within-catch-up in HAZ, especially before age 5
F2014	Ethiopia, India, Peru, Vietnam	–	✓	–	✓	Some within-catch-up in HAZ, no between catch-up
L2014	Brazil, Guatemala, India, Philippines, RSA	✓	✓	–	–	Some between catch-up in HAZ, none in HAD
L2015	Ethiopia, India, Peru, Vietnam	✓	✓	–	–	No between catch-up in HAZ
T2015	Malawi	✓	✓	–	–	Some between catch-up in HAZ, none in HAD
H2016	China, South Africa, Nicaragua	–	–	–	✓	Incomplete within catch-up in HAZ
Z2016	Bolivia (Amazon)	–	✓	✓	✓	Incomplete within catch-up in HAZ, persistent height deficits
S2017	Timor-Leste	–	✓	–	–	Some between catch-up in HAZ, not in BMI (BAZ)

Notes: Studies are classified by us according to the main form(s) of the catch-up parameter on which each study focuses (between-versus within-group), as well the outcome metrics used (raw, e.g., HAD score; versus relative, e.g., HAZ score); studies were identified from Google Scholar in December 2017 based on the combined search text: ' "catch-up growth" & children & "developing countries" '. The search period was limited to the period 2012–2017 and only primary empirical studies undertaken on developing countries were retained; also, studies concerned with catch-up growth either after specific interventions or associated with specific medical conditions were excluded.
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. (2018).

of the studies in Table 1 include an analysis of recovery from stunting, which is identified via a binary transformation of HAZ scores. 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 appears 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 two different forms of catch-up, their findings are rarely systematized so as to shed light on the correspondence between alternative forms of convergence. More generally, evidence of a fall in the prevalence of stunting in a sample (group) over time may be consistent with either between- or within-group catch-up (or both). That is, without looking at trends in the rest of the distribution, we cannot say whether the reduction was associated with improvements across the sample (e.g., between-group catch-up) as opposed to more rapid growth of the initially-stunted, possibly in the context of a stable or even worsening of the sample mean. This further motivates the importance of the conceptual distinctions sketched above.

3. Empirical methods

The previous section distinguished between various types of catch-up growth, noting that previous empirical studies have often considered particular definitions but neglected others. One reason for the limited scope of past studies may be methodological. Conventional approaches to testing for catch-up, such as based on a dynamic panel (longitudinal) specification, generally permit only specific forms of catch-up to be investigated. Without reviewing these approaches in detail here due to space limitations, in this section we show how a latent growth model provides a unified testing framework.

To start, a general latent growth model for child development can be expressed as follows:

$$y_i(t_i) = (\alpha_0 + \alpha_i) + (\beta_0 + \beta_i)f(t_i) + x_{it}'\gamma + \varepsilon_{it} \tag{1a}$$

$$\mathrm{Var}(\alpha_i) = \sigma_\alpha^2,\ \ \mathrm{Var}(\beta_i) = \sigma_\beta^2, \mathrm{E}(\alpha_i\beta_i|x_{it}) = \rho_{\alpha\beta}(\sigma_\alpha\sigma_\beta) \tag{1b}$$

$$\mathrm{E}(x_{it}'\alpha_i|\beta_i) = \rho_{x\alpha}(\sigma_\alpha\sigma_x), \mathrm{E}(x_{it}'\beta_i|\alpha_i) = \rho_{x\beta}(\sigma_x\sigma_\beta) \tag{1c}$$

where y is an anthropometric outcome of interest; i indexes individual children, whom we assume are observed on more than one occasion; t represents the child's age at the time of measurement; x is a vector of observed explanatory variables (e.g., household wealth); and ε is residual error (e.g., due to measurement). Consistent with our interest in testing for both between- and within-group catch-up, we assume henceforth that the outcome y is referenced to an external population. The nature and interpretation of the intercept (α) and slope (β) coefficients are important. Both are comprised of two parts: first is the sampled group average, taking a zero subscript; and, second, unobserved individual-specific components, taking the i subscript, which are centered on zero by construction. The α terms capture the expected level of the outcome variable (conditional on covariates) when $f(t_i) = 0$; and where $f(\cdot)$ is an unspecified functional form. The β terms represent the rate or velocity of growth in the outcome with age. The two unobserved components correspond to estimates of deviations from mean child size and growth velocity respectively (using the terminology of Tanner, 1962).

Focusing on the metrics of catch-up growth, estimates for the average slope β_0 capture the mean velocity of growth in the sample. Thus, as per externally-referenced HAZ or HAD scores, positive estimates indicate a faster average pace of growth compared 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 (between-group) catch-up has occurred. Within-group catch-up 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 with age. Thus, within-population catch-up requires $\rho_{\alpha\beta} < 0$, where the latter correlation is conditional on the included controls (x_{it}).

Based on this framework, Table 2 sets out specific hypotheses regarding the combinations of catch-up growth that may be identified from the model. Formally, convergence of the sample mean toward the median of a healthy external reference population requires: $\alpha_0 < 0 \wedge \beta_0 > 0$; while within catch-up requires $\rho_{\alpha\beta} < 0$. The appropriate null hypotheses to be taken to the data, corresponding to an absence of either form of catch-up, are 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 one-sided statistical tests are appropriate here. Put differently, conventional tests of a general null hypothesis that, say, $\beta_0 \neq 0$ are not

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

		Between convergence?	
		Yes	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: '∧' is the logical AND operator; '∨' 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.

particularly informative about the presence of (positive) catch-up growth.

The latent growth framework holds a number of advantages in comparison to more conventional dynamic panel specifications, where catch-up growth is investigated by looking at the extent to which outcomes (e.g., height) in a previous period predict current outcomes. First, and as indicated above, the framework permits simultaneous but distinct estimation of *both* within and between forms of catch-up growth. Second, equation (1a) contains no lagged outcome(s) in the set of explanatory variables. Thus, measurement error in the outcome variable is not expected to bias estimates of the catch-up parameters and no instrumental variables procedure is required to address this (e.g., Mani, 2012). Third, the absence of a lagged outcome variable means the unobserved individual intercepts and slopes can be retrieved easily and directly (see further below), avoiding technical problems of doing so using dynamic panel estimators (Bond et al., 2005). Fourth, the framework extends to multiple observation periods and theoretically permits non-linearities in growth patterns to be estimated and tested (Chirwa et al., 2014).

A latent growth framework does have some weaknesses. As with any regression model, there is scope for bias from measurement and specification error. The former relates particularly to child age (which may be rounded down or subject to recall mistakes). The latter particularly refers to the functional form for child age. This is critical; but since we have only four observations per individual, in our application (Section 4) non-linear functional forms cannot be considered. Thus, at best, our results can be interpreted as a (possibly, rough) linear approximation to the true underlying process. A related concern is the implicit assumption that the evolution of child height is a trend stationary process. If not, then this class of models would be inappropriate. Either way, shocks to growth are likely to be highly persistent implying that attention needs to be given to the error structure, perhaps even necessitating inclusion of autoregressive terms (Hamaker, 2005). Thus, we believe further research is required to test the performance of alternative models and estimation procedures for child growth.

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 samples from communities in four low-income countries. Since the YL data have been used extensively before, detailed introduction is not necessary (e.g., see Barnett et al., 2012, Crookston et al., 2013, Lundeen et al., 2014a). The Young Lives study received institutional review board approval from the University of Oxford and local ethics review boards at each participating country's lead institution. The study has been performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments. Informed consent was obtained from all individual participants involved in the study. For this article, only de-identified public-use data are used so no further ethical approval has been sought.

Section 4.1 briefly introduces the data; Section 4.2 reports the baseline results for tests of the catch-up hypotheses (as outlined in Table 2); and Appendix B explores time-invariant child characteristics that are associated with systematic differences in the size and direction of catch-up parameters. Throughout this section we analyze catch-up in terms of children's height-for-age, as measured on a continuous scale. This is consistent with the focus of many existing studies, the rationale being that low height-for-age is understood to be a key marker of chronic under-nutrition and appears to be robustly associated with adverse outcomes in later life (Hoddinott et al., 2013; Almond et al., 2017). Also, since conventional measures of stunting are based on binary transformations of HAZ scores, we retain more information using the HAZ score itself.

4.1. Data

We use the four latest public-use rounds of the YL data and focus exclusively on the younger cohort, who were approximately one year old in the first round (collected in 2002) and around 12 years old in the fourth round (collected in 2013). The older cohorts are excluded since academic analysis of growth recovery (faltering) has typically focussed on the growth trajectories of younger children (e.g., Victora et al., 2010; Leroy et al., 2014). Table 3 summarizes the analytical dataset we use hereafter. We have removed children who have missing data for their age (1116 observations removed), or who were under 6 months old in the first round (165 observations removed); and we set outlier observations on height outcomes to missing (371 cases, defined as observations with HAZ scores of ± 5 standard deviations). Specifically, heights corresponding to HAZ scores of ± 5 (standard deviations) are set to missing. Also, children with a median of absolute changes in HAZ scores that is greater than 4 standard deviations are treated as missing. 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 30,373 observations on 7641 children. As Part (a) of the table shows, the data covers almost 2000 children in each of the four countries. It 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.

The descriptive statistics in Table 3(a) report averages across rounds for various height outcomes. 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. The various metrics we report are: (i) raw height (in cms); (ii) the HAZ score (relative measure); (ii) the raw height difference (HAD, in cms); and (iv) measures of stunting, derived from the HAZ scores. Part (b) of the table reports sample averages of long differences between the final and first rounds (calculated in pairwise fashion over each child) for the same metrics. As expected, all countries display significant gains in mean height, ranging from gains of 68–72 cms over around 11 years. However, the other height metrics, which are all referenced to the WHO's 2007 child growth charts (see Vidmar et al., 2013), paint a diverse picture. The raw differences (HAD) indicate a deterioration in outcomes over time relative to the reference population in all four countries. In contrast, changes in relative scores (HAZ) are in the positive domain in all countries, other than India. The metrics of stunting generally also

Table 3
Descriptive statistics, by country.
Source: own estimates.

	Ethiopia		India		Peru		Vietnam	
	Mean	(st.err.)	Mean	(st.err.)	Mean	(st.err.)	Mean	(st.err.)
(a) Averages (Rounds 1 to 4):								
No. children	1866	(0.06)	1917	(0.10)	1900	(0.29)	1911	(0.22)
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)	49.9	(0.57)	48.7	(0.57)
Household size	6.0	(0.02)	5.3	(0.03)	5.5	(0.02)	4.7	(0.02)
Wealth index	−1.0	(0.01)	0.1	(0.01)	0.3	(0.02)	0.6	(0.01)
Mom's age at birth	26.4	(0.08)	22.6	(0.05)	25.8	(0.08)	26.1	(0.07)
Mom's years schooling	2.7	(0.05)	3.7	(0.05)	7.6	(0.05)	6.9	(0.04)
Height	109.6	(0.30)	108.8	(0.29)	109.4	(0.31)	110.6	(0.31)
Height-for-age *z*-score (HAZ)	−1.4	(0.01)	−1.4	(0.01)	−1.3	(0.01)	−1.1	(0.01)
Height-for-age diff. (HAD)	−6.9	(0.07)	−7.3	(0.07)	−6.1	(0.07)	−5.6	(0.07)
Not stunted (%)	70.6	(0.53)	69.6	(0.53)	74.6	(0.50)	79.1	(0.47)
Not severely stunted (%)	90.8	(0.34)	93.2	(0.29)	93.7	(0.28)	96.0	(0.23)
(b) Long differences (Round 4 - Round 1):								
No. children	−13	(0.00)	−21	(0.00)	−66	(0.00)	−44	(0.00)
Child's age (years)	11.1	(0.01)	11.0	(0.01)	10.9	(0.01)	11.2	(0.01)
Household size	0.1	(0.06)	−0.6	(0.06)	−0.5	(0.06)	−0.4	(0.04)
Wealth index	1.0	(0.02)	1.2	(0.02)	1.1	(0.02)	1.2	(0.02)
Height	69.5	(0.16)	68.4	(0.16)	71.2	(0.15)	71.9	(0.17)
Height-for-age *z*-score (HAZ)	0.0	(0.04)	−0.2	(0.03)	0.4	(0.02)	0.0	(0.02)
Height-for-age diff. (HAD)	−6.4	(0.15)	−6.6	(0.14)	−3.5	(0.15)	−4.5	(0.16)
Not stunted (%)	10.3	(1.31)	1.0	(1.22)	11.4	(1.06)	0.4	(0.99)
Not severely stunted (%)	12.8	(1.01)	4.3	(0.74)	6.1	(0.70)	−0.3	(0.57)

Note: aside from raw height, all anthropometric outcomes are calculated using the WHO's 2007 reference charts; stunting is given by a HAZ score below −2; severe stunting is given by a HAZ score below −3; see text for further description of height transformations.

show some moderate improvements over time, but in Vietnam and India this is dependent on the measure of stunting employed.

Fig. 1 illustrates trends in the HAZ and HAD metrics for the four countries over the four rounds. Each figure plots the estimated linear trend (by age) for the 5th percentile, the median and the 95th percentile of the relevant score distributions. The figures support the overall pattern in aggregate outcomes described in Table 3(b). Additionally, they reveal key differences in how the shapes of the distributions have evolved. The HAZ plots (Fig. 1a) indicate the distributions for Ethiopia and India have narrowed over time (with age), which would be consistent with a process of within-group convergence in the *z*-scores. But the same plots reveal no clear change in the median score of the same samples, thereby confirming the relevance of distinguishing between within- and between-group catch-up processes. At the same time, the corresponding HAD plots (Fig. 1b) point to distributional divergence in all cases, as well as deterioration or divergence of the sample median from the reference median. In line with a number of studies discussed in Table 1, this would support the contention that conclusions about catch-up are sensitive to whether outcomes are expressed in raw or relative terms.

4.2. Baseline results

This sub-section summarizes our estimates of the latent growth model. Intuitively, this amounts to modelling the trends illustrated in Fig. 1, both for the central tendencies and the shapes of the distributions. Table 4 reports estimates based on equation (1a) for the four YL countries, comparing both the relative (HAZ) and raw (HAD) outcomes. Although we use different height outcomes, the same right-side specification is employed throughout. The included control variables align with those used in previous studies; but there are some limitations as we can only retain variables that are common to all countries and rounds of the YL datasets. (for further discussion see the Supplementary Material). Specifically, the model includes: child age (t, expressed in years), entered in de-meaned linear form; and a set of time-varying control variables (x), which adjust for short-run variation in outcomes and provide more precise identification of the (fixed) latent components. The elements of x are: a household wealth index (based on the standardized first principal component of assets owned by each household); household size; the number of siblings in different age groups; whether the mother is the primary care-giver; and two metrics of household exposure to shocks (based on principal components scores).

For estimation, we use a fixed effects estimator with individual slopes (FEIS), implemented in Stata v15.1 via the user-written command reghdfe (Correia, 2017). Appendix A discusses the properties of this estimator, which is found to outperform alternatives such as a (correlated) mixed effects estimator. The individual-level latent effects (α_i, β_i) are retrieved after estimation (see Guimarães and Portugal, 2010); and the correlation coefficient ($\rho_{\alpha\beta}$) is calculated directly from these estimated fixed effects. Of course, by construction of equation (1a), all time-invariant factors and their interaction with the child's age are captured by the latent variables. Therefore, differences in latent growth patterns along fixed characteristics, such as gender (or location), cannot be seen directly from the regression estimates. Nonetheless any such relationships can be revealed by secondary or stratified analysis of the estimated latent factors. By way of example, Appendix Tables C1–C2 replicate Table 4 (using the same regression results), showing estimates for boys and girls separately; and Appendix B extends this to a multivariate setting.

For each country and outcome, the first three columns of Table 4 report estimates of the three regression 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 ($\rho_{\alpha\beta}$). Following Table 2, these provide the basis for evaluating what combination of catch-up (between/within) obtains for the given outcome and country. Specifically, for each parameter we report the probability associated with the relevant one-sided null hypothesis (denoted H0). As described further in Appendix A we combine the individual tests to

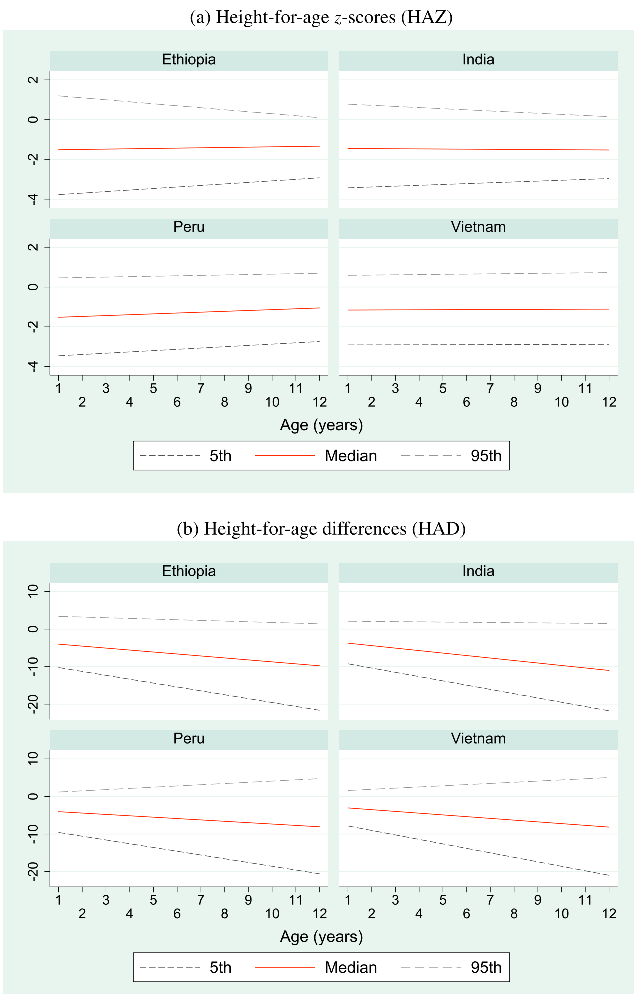

Fig. 1. Growth patterns by country.
Notes: trends estimated via quantile regressions.
Source: own estimates.

determine whether either between-group, within-group or both forms of catch-up are found. These conclusions are summarized in the '?' columns, where a N(o) indicates 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. Linking to the four-way typology set out earlier, Appendix Table C3 summarizes the same results in the spirit of Table 1.

Four main findings stand out. First, as reasonably expected given the nature of the YL surveys, estimates for mean stature ($\hat{\alpha}_0$) are below that of the reference population in all four countries (for both HAZ and HAD scores). 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. Second, 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. Average HAZ scores in the Peruvian and Vietnamese samples are growing significantly faster than in the reference population, meaning we reject the null hypothesis that $\beta_0 \leq 0$. However, in Ethiopia and India we find no evidence of above-normal growth rates in HAZ scores. Thus, from this perspective, relative between-group catch-up has taken place in Peru and Vietnam, but not elsewhere.

Third, evidence for relative within-group catch-up is somewhat mixed. The estimates for $\rho_{\alpha\beta}$ using the HAZ scores are negative and significant in all countries other than Vietnam. A negative coefficient indicates that shorter (taller) children are growing faster (slower) on

Table 4
Estimates of catch-up growth, by country and type.
Source: own estimates.

	α_0	β_0	$\rho_{\alpha\beta}$	Between			Within		Both
				$H0_{\alpha_0}$	$H0_{\beta_0}$	?	$H0_{\rho_{\alpha\beta}}$	?	?
(a) HAZ scores (relative outcome):									
Ethiopia	−1.387	0.003	−0.298	0.00	0.32	N	0.00	Y	N
	(0.060)	(0.007)	(0.040)						
India	−1.444	−0.014	−0.190	0.00	0.98	N	0.00	Y	N
	(0.045)	(0.007)	(0.053)						
Peru	−1.259	0.048	−0.150	0.00	0.00	Y	0.00	Y	Y
	(0.092)	(0.004)	(0.035)						
Vietnam	−1.126	0.020	0.035	0.00	0.00	Y	0.82	N	N
	(0.086)	(0.004)	(0.039)						
(b) HAD scores (raw outcome):									
Ethiopia	−6.916	−0.519	0.528	0.00	0.99	N	0.99	N	N
	(0.301)	(0.034)	(0.037)						
India	−7.316	−0.613	0.632	0.00	0.99	N	0.99	N	N
	(0.217)	(0.031)	(0.036)						
Peru	−6.637	−0.167	0.357	0.00	0.99	N	0.99	N	N
	(0.495)	(0.030)	(0.036)						
Vietnam	−5.614	−0.286	0.762	0.00	0.99	N	0.99	N	N
	(0.453)	(0.041)	(0.021)						

Note: the first three columns report parameter estimates (and cluster-robust standard errors, in parentheses) from the latent growth specification, including 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.

average than their counterparts. It follows there is somewhat more consistent evidence of within-group catch-up in HAZ scores in the YL countries (i.e. in three of the four countries). However, as indicated in Appendix Table C3, there is no systematic correspondence across the two types of catch-up – only Peru displays evidence of *both* between- and within-group catch-up simultaneously.

Fourth, in clear contrast to the evidence regarding relative types of catch-up, Table 4(b) indicates there is no evidence of catch-up growth when raw (HAD) scores are employed. In fact, and in line with Fig. 1b, the HAD estimates for both β_0 and $\rho_{\alpha\beta}$ point to materially divergent growth paths – i.e., the sample mean is falling behind the reference mean and, within the sample, taller children are extending their raw height advantage over time compared to their peers in the same sample. This underlines the markedly different conclusions that emerge about catch-up growth depending on how growth outcomes are expressed.

Finally, while the above tests point to at least one type of relative catch-up growth in all countries, the point estimates in Table 4(a) indicate the pace of between- and within-group catch-up is (at best) moderate and incomplete. For instance, Peru displays the largest point estimate for β_0, but it would take around 27 years (on the same trajectory) for the mean HAZ score in this sample to converge to zero. Fig. 2 combines these 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, taking into account the estimated correlation with the slope effect. The results confirm the slow pace of between-group catch-up – i.e., over the 11 years spanned by the data, the mean (predicted) HAZ score for Peru and Vietnam increased by less than 0.5 of a standard deviation. Similarly, for those countries displaying within-country catch-up, this 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 and 12; and these two trends would take about 30 years to converge.

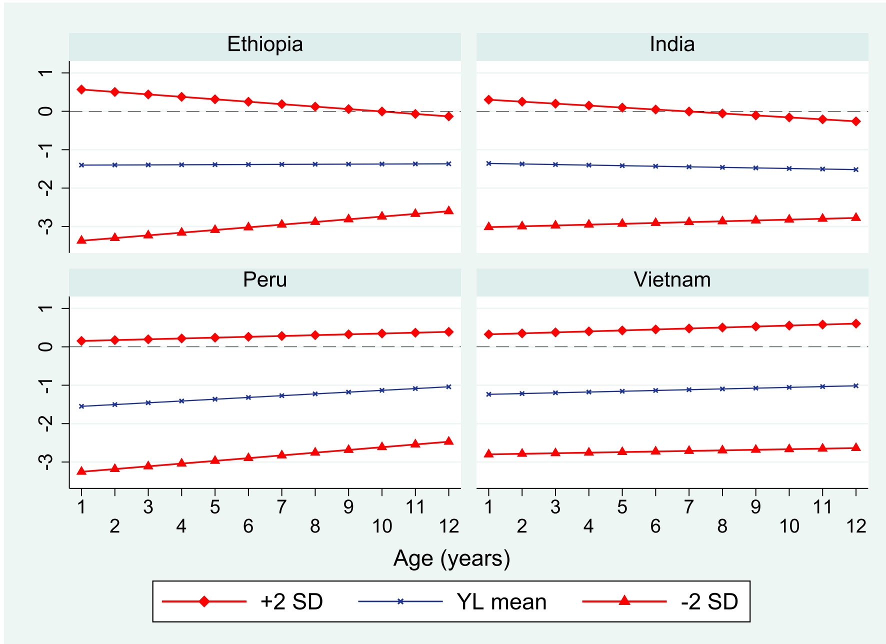

Fig. 2. Estimated HAZ score trajectories, by country.
Note: trends are derived from the estimates reported in Table 4(a); the blue line (YL mean) is the estimate of the mean HAZ score for the sample; the two red lines are estimated at ± two standard deviations (SDs) of the estimated individual fixed-effects distributions; dashed horizontal line at $y = 0$ is the reference median.
Source: own estimates.

5. Discussion and conclusion

The first aim of this study was to revisit how catch-up growth in children is defined. Our motivation was that existing definitions tend to focus on specific aspects of catch-up, such as either between-group or within-group convergence, but not both. We also argued that existing studies tend to focus on specific outcomes (e.g., relative versus raw deviations from medians). Consequently, there is limited comparative evidence regarding the range of different types of catch-up and how they relate to one another. Addressing this concern, we set out a four-way classification of types of catch-up and argued that a latent growth framework provides an attractive unified setting for empirical testing. We also 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 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 investigated patterns of catch-up in the 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 stature than those from a healthy international reference population. On the basis of externally-standardized HAZ scores (relative outcomes), 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 in HAZ scores, but again this appears to have proceeded at a relatively moderate pace. Together, and as depicted in Fig. 2, 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. However, when we shift to raw HAD scores, there is no evidence of catch-up in any of the samples.

The results clearly demonstrated that estimates of both forms of catch-up are highly sensitive to how outcomes are expressed. Differences between outcome metrics reflect a range of factors, including (*inter alia*): their sensitivity to changes in the shape of the reference distribution; how initial gaps compound over time; and their focus on particular aspects of the outcome distribution (e.g., stunting). Since the construction choices underlying specific outcomes often correspond to different substantive research interests, there is unlikely to be a single 'best' metric for all situations. At the same time, we note that at any fixed age, the HAZ and HAD scores will be perfectly correlated and therefore can be used equally well as a predictor for other outcomes (e.g., cognitive skills). Furthermore, a reliance on HAD scores (implicitly) suggests that a given height deficit relative to the median height in a well-nourished population is equally meaningful, regardless of the actual level of this reference median. This is a very strong assumption and (as we have seen) indicates that such absolute differences are likely to yield a hard lower bound to the extent of catch-up growth.

In the Supplementary Material we extended the analysis to show how the latent growth framework can be used to identify systematic drivers of heterogeneity in key parameters of individual growth trajectories – child mean size, growth velocity and their correlation. This was developed in a second analytical stage, whereby 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 schooling attainment, as well as community fixed effects. These results revealed both similarities and differences in patterns across the four countries. Higher household wealth is often associated with more rapid growth, enabling children to extend initial advantages. And, community fixed effects appear to play a material role in influencing growth trajectories. However, differences in gender and treatment of first-borns appear to vary across the samples. The analysis therefore reveals systematic differences in the forms of catch-up growth and their underlying drivers across different disadvantaged communities in low- and middle-income countries. These suggest that child developmental interventions need to be tailored carefully to each context.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.socscimed.2018.07.003.

References

Adair, L.S., 1999. Filipino children exhibit catch-up growth from age 2 to 12 years. J. Nutr. 129 (6), 1140–1148.

Almond, D., Currie, J., Duque, V., 2017. Childhood Circumstances And Adult Outcomes: act II. Journal of Economic Literature forthcoming.

Anand, P., Roope, L., 2016. The development and happiness of very young children. Soc. Choice Welfare 47 (4), 825–851.

Barnett, I., Ariana, P., Petrou, S., Penny, M.E., Le, T.D., Galab, S., Woldehanna, T., Escobal, J.A., Plugge, E., Boyden, J., 2012. Cohort profile: the young lives study. Int. J. Epidemiol. 42 (3), 701–708.

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. Lancet 389 (10064), 77–90.

Bond, S., Nauges, C., 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. https://ideas.repec.org/p/ifs/cemmap/07-05.html.

Britto, P.R., Lye, S.J., Proulx, K., Yousafzai, A.K., Matthews, S.G., Vaivada, T., Perez-Escamilla, R., Rao, N., Ip, P., Fernald, L.C.H., MacMillan, H., Hanson, M., Wachs, T.D., Yao, H., Yoshikawa, H., Cerezo, A., Leckman, J.F., Bhutta, Z.A., 2017. Nurturing care: promoting early childhood development. Lancet 389 (10064), 91–102.

Cameron, N., Preece, M.A., Cole, T.J., 2005. Catch-up growth or regression to the mean? Recovery from stunting revisited. Am. J. Hum. Biol. 17 (4), 412–417.

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

Coast, J., Smith, R., Lorgelly, P., 2008. Should the capability approach be applied in health economics? Health Econ. 17 (6), 667–670.

Correia, S., 2017. REGHDFE: Stata module for linear and instrumental-variable/GMM regression absorbing multiple levels of fixed effects. Technical Report Statistical Software Components s457874. Boston College Department of Economics. https://EconPapers.repec.org/RePEc:boc:bocode:s457874.

Crookston, B.T., Schott, W., Cueto, S., Dearden, K.A., Engle, P., Georgiadis, A., Lundeen, E.A., Penny, M.E., Stein, A.D., Behrman, J.R., 2013. Postinfancy growth, schooling, and cognitive achievement: young lives. Am. J. Clin. Nutr. 98 (6), 1555–1563.

Desmond, C., Casale, D., 2017. Catch-up growth in stunted children: definitions and predictors. PLoS One 12 (12), e0189135.

Fink, G., Rockers, P.C., 2014. Childhood growth, schooling, and cognitive development: further evidence from the Young Lives study. Am. J. Clin. Nutr. 100 (1), 182–188.

Guimarães, P., Portugal, P., 2010. A simple feasible procedure to fit models with high-dimensional fixed effects. STATA J. 10 (4), 628–649.

Haisma, H., Yousefzadeh, S., Boele Van Hensbroek, P., 2017. Towards a capability approach to child growth: a theoretical framework. Matern. Child Nutr. 14 (2), e12534.

Hamaker, E.L., 2005. Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. Socio. Meth. Res. 33 (3), 404–416.

Handa, S., Peterman, A., 2016. Is there catch-up growth? Evidence from three continents. Oxf. Bull. Econ. Stat. 78 (4), 470–500.

Heckman, J.J., 2007. The economics, technology, and neuroscience of human capability formation. Proc. Natl. Acad. Sci. Unit. States Am. 104 (33), 13250–13255.

Hirvonen, K., 2014. Measuring catch-up growth in malnourished populations. Ann. Hum. Biol. 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., Martorell, R., 2013. Adult consequences of growth failure in early childhood. Am. J. Clin. Nutr. 98 (5), 1170.

Jones, S., Behrman, J.R., Dang, H.H., Anand, P., 2018. Identifying Catch-Up Trajectories In Child Growth: New Methods With Evidence From Young Lives. Policy Research Working Paper 8353. The World Bank.

Leroy, J.L., Ruel, M., Habicht, J.P., 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. J. Nutr. 144 (9), 1460–1466.

Leroy, J.L., Ruel, M., Habicht, J.P., Frongillo, E.A., 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 Pediatr. 15 (1), 145.

Lundeen, E.A., Behrman, J.R., Crookston, B.T., Dearden, K.A., Engle, P., Georgiadis, A., Penny, M.E., Stein, A.D., 2014a. Growth faltering and recovery in children aged 1–8 years in four low-and middle-income countries: young Lives. Publ. Health Nutr. 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. Am. J. Clin. Nutr. 100 (3), 821–825.

Mani, S., 2012. Is there complete, partial, or no recovery from childhood malnutrition? empirical evidence from Indonesia. Oxf. Bull. Econ. Stat. 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. J. Epidemiol. Community Health 71 (10), 981.

de Onis, M., Blössner, M., Borghi, E., 2012. Prevalence and trends of stunting among preschool children, 1990–2020. Publ. Health Nutr. 15 (1), 142–148.

Outes, I., Porter, C., 2013. Catching up from early nutritional deficits? Evidence from rural Ethiopia. Econ. Hum. Biol. 11 (2), 148–163.

Richter, L.M., Daelmans, B., Lombardi, J., Heymann, J., Boo, F.L., Behrman, J.R., Lu, C., Lucas, J.E., Perez-Escamilla, R., Dua, T., Bhutta, Z.A., Stenberg, K., Gertler, P., Darmstadt, G.L., 2017. Investing in the foundation of sustainable development: pathways to scale up for early childhood development. Lancet 389 (10064), 103–118.

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. Soc. Sci. Med. 97 (2013), 278–287.

Sen, A., 2002. Why health equity? J. Health Econ. 11, 659–666.

Spencer, P.R., Sanders, K.A., Judge, D.S., February 2018. Growth curves and the international standard: how children's growth reflects challenging conditions in rural Timor-Leste. Am. J. Phys. Anthropol. 165 (2), 286–298.

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

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

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

Wit, J.M., Boersma, B., 2002. Catch-up growth: definition, mechanisms, and models. J. Pediatr. Endocrinol. Metabol. 15, 1229–1242.

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