Empir Econ
https://doi.org/10.1007/s00181-018-1421-z

# The origins of early childhood anthropometric persistence

**Augustine Denteh**[1] · **Daniel L. Millimet**[2,3] · **Rusty Tchernis**[1,3,4]

Received: 14 November 2017 / Accepted: 22 January 2018


**Abstract** The rates of childhood obesity have increased dramatically in the last few decades. Non-causal evidence suggests that childhood obesity is highly persistent over the life cycle. However, little is known about the origins of this persistence. This paper examines the evolution of anthropometric measures from birth through primary school. We provide estimates of the causal effect of past anthropometric outcomes on future anthropometric outcomes (state dependence) and investigate the importance of time-varying and time-invariant factors in the dynamics of childhood anthropometric measures. We find that anthropometric measures are highly persistent from infancy through primary school. Moreover, most of this persistence is driven by time-invariant, unobserved factors that are determined prior to birth, consistent with the so-called fetal origins hypothesis. Thus, policy interventions designed to improve childhood obesity will only have meaningful, long-run effects if these time-invariant, unobserved factors are altered. Future research is needed to identify such factors, although evidence suggests that maternal nutrition may play an important role.

This study was conducted by Georgia State University and Southern Methodist University under a cooperative agreement with the U.S. Department of Agriculture, Economic Research Service, Food and Nutrition Assistance Research Program (Agreement No. 58-5000-0-0080). The views expressed here are those of the authors and do not necessarily reflect those of the USDA or ERS. The authors are grateful to Chris Ruhm for helpful comments and Lorenzo Almada for excellent research assistance.

✉ Rusty Tchernis
rtchernis@gsu.edu

1 Department of Economics, Andrew Young School of Policy Studies, Georgia State University, P. O. Box 3992, Atlanta, GA 30302-3992, USA

2 Department of Economics, Southern Methodist University, Dallas, USA

3 IZA, Bonn, Germany

4 NBER, Cambridge, MA, USA

Published online: 25 May 2018

**Keywords** Childhood obesity · Persistence · Fetal origins hypothesis

**JEL Classification** C23 · I12 · I18

## 1 Introduction

The rise in childhood obesity in the USA is well chronicled (Ogden et al. 2002, 2008; Skinner et al. 2016). Moreover, non-causal evidence suggests that childhood obesity is highly persistent over the life cycle (Serdula et al. 1993; Power et al. 1997). However, little is known about the origins of this persistence. In this paper we attempt to answer three questions. First, how do weight, height, and body mass index (BMI) evolve from birth through primary school? Second, what is the *causal* effect of past weight and height status on the future weight and height of children? In other words, how important is *state dependence* in the evolution of anthropometric measures during the early part of the life cycle. Third, how important are time-varying and time-invariant factors in the dynamics of childhood anthropometric measures? Our analysis is fundamentally important for researchers and policy-makers involved in formulating policies and interventions to address the obesity epidemic. For instance, if obesity originates early in life and is persistent over time, then early childhood intervention is preferable to later in adolescence or beyond.[1] However, and perhaps more importantly, if persistence is due to time-invariant underlying factors rather than state dependence, then only by altering these factors can children be moved onto a different trajectory.

The prevalence of obese adolescents has tripled over the last thirty years; it has more than doubled for younger children. Defined as having an age- and sex-adjusted BMI above the 95th percentile of the reference distribution, the prevalence of obese children increased from 5 to 12.4% for 2–5-year-old children and from 5 to 17.6% for 12–19-year-olds between 1976 and 2006 (Ogden et al. 2008). In addition, vast differences in the time trends of BMI increases have been documented: The incidence of obesity among white girls aged 12–19 has increased from 7.4 to 14.5% between 1988 and 2006, whereas the corresponding figures for African-American girls are 13.2 and 27.7% (Ogden et al. 2002, 2008). Epidemic increases in childhood obesity have also been observed among preschool children (Deckelbaum and Williams 2001, p. 242S).

While there exists some evidence that childhood obesity rates may have begun to recede or at least stabilized in the USA, public health concerns over childhood obesity remain high due to the well-documented consequences of obesity and the lack of understanding pertaining to the turnaround (Ogden et al. 2014, 2016). Obesity burdens individuals with severe physical, economic, and emotional suffering and puts children and adolescents at risk of a number of health problems such as those affecting cardiovascular health, the endocrine system, and mental health (Deckelbaum and Williams

[1] For instance, Eriksson et al. (2001, p. 735) conclude that "obesity is initiated early in life." Also, while the evidence is not ironclad, there is suggestive evidence that children may become entrenched on an "obesity trajectory" even before kindergarten (Taveras et al. 2008, 2010; Lamb et al. 2010; Gillman et al. 2008). While public health officials tend to advocate school-based reforms in light of the near-universal enrollment, others stress the importance of preschool interventions [e.g., Frisvold and Giri (2014), Dietz and Gortmaker (2001), Davis and Christoffel (1994)].

2001; Krebs et al. 2003). For instance, Dietz and Gortmaker (2001) note that 60% of overweight children aged five to ten years have at least one associated cardiovascular disease risk factor.

Perhaps the most crippling consequence of childhood obesity stems from its impact on adult obesity (Serdula et al. 1993; Power et al. 1997), with its concomitant effects. Currently, 60% of the total US population is overweight or obese and 50% is expected to be obese in 2030 at the current rate (Dor et al. 2010). Walpole et al. (2012) calculate that, due to obesity, North America accounts for 34% of the total human biomass in the world despite containing only 6% of the world population. Moreover, the authors estimate that, if the entire world had the same age-sex BMI distribution as the USA, total human biomass will increase by 58 million tonnes which would be equivalent to 935 million people of world average BMI. Finkelstein and Zuckerman (2010) report that if the childhood obesity epidemic continues unabated at the current rate, as many as 30–40% of the US population will develop type 2 diabetes during their lifetime. Mocan and Tekin (2011) document the links between adult obesity and lower wages, productivity, and self-esteem. In the USA, the total cost attributable to obesity was over $75 billion in 2000 according to Finkelstein et al. (2004); more recent estimates put the cost over $200 billion (Cawley and Meyerhoefer 2012).

While the changes in childhood obesity rates across cohorts, as well as the consequences of these increases, are well documented, much less is known about how anthropometric measures of children evolve over the life cycle. A growing literature has investigated persistence in anthropometric outcomes in a non-causal framework, stressing the correlation in outcomes over time. [2] Whitaker et al. (1997) found that the probability of an overweight six-year-old child becoming an obese adult is 50% compared to 10% for a non-overweight child. In addition, the risk of becoming obese in adulthood is exacerbated by having an obese parent (Whitaker et al. 1997). Eriksson et al. (2001) found that individuals with BMI greater than 16 as opposed to below 14.5 at age seven were three times more likely to be obese as an adult. Nader et al. (2006) find that children who were overweight prior to the age of 5 are five times as likely to be overweight at 12 relative to children who were not overweight prior to the age of 5.

Freedman et al. (2001) also report a strong relationship between overweight status in childhood and adult BMI. However, they find a striking result that obese adults who were overweight prior to age eight have a much higher BMI than individuals suffering from adult-onset obesity (41 vs. 35). Van Cleave et al. (2010) analyze changes in the prevalence of obesity and other chronic conditions (e.g., asthma, other physical and learning conditions). The authors find that the prevalence of obesity is increasing and is highly persistent over time. Conversely, many children with chronic conditions at ages two through eight did not have the condition six years later. Finally, Millimet and Tchernis (2013a) assess the persistence of anthropometric outcomes during infancy and primary school, documenting a significant increase in persistence upon entry into primary school. Deckelbaum and Williams (2001, p. 239S) conclude: "Disturbingly, obesity in childhood, particularly in adolescence is a key predictor for obesity in

[2] Iughetti et al. (2008) provide an excellent review.

adulthood." Similarly, Dietz and Gortmaker (2001, p. 340) state: "The best evidence suggests that the majority of overweight adolescents go on to be overweight adults."

In addition to the literature already discussed pertaining to the correlations between childhood weight status and adult obesity, a related literature investigates the persistence in health among adolescents and adults in a *causal* framework. For example, Halliday (2008) investigates persistence in self-reported health status among white adults aged 22–60 using data from the Panel Study of Income Dynamics (PSID) and allows the parameters of the model to vary. The results suggest that the degree of state dependence—the causal effect of past states on one's current state—in health is modest for half the population, yet it explains much of the observed persistence in health for the other half. Ham et al. (2013) analyze persistence in bulimia nervosa in young women. The authors find a substantial role for state dependence in the persistence of bulimia nervosa. Our analysis follows the logic of these studies.

This paper contributes to the existing literature by attempting to understand the origins of the persistence in childhood anthropometric measures. We revisit the question of persistence in early childhood anthropometric outcomes, investigating the relative importance of state dependence (i.e., a causal effect of past anthropometric status on future anthropometric status), unobserved heterogeneity (i.e., unobserved genetic or environmental risk factors), and observed heterogeneity (i.e., commonly measured risk factors) on this persistence. We do so by estimating dynamic regression models using data from two samples: the Early Childhood Longitudinal Survey—Kindergarten Cohort (ECLS-K) and the Early Childhood Longitudinal Survey—Birth Cohort (ECLS-B). The ECLS-K is a nationally representative longitudinal survey of children entering kindergarten in Fall 1998. In addition to providing information on birthweight, anthropometric data are collected at several points in time between kindergarten and eighth grade. The ECLS-B is also a nationally representative longitudinal survey of children born in the USA in 2001. Information is provided on these children at ages 9 months, two years, four years, and five years. These two samples are complementary because while the ECLS-K sample permits an examination of anthropometric persistence from kindergarten through primary school, the ECLS-B sample allows for a more refined examination of anthropometric trajectories prior to kindergarten entry.

Two salient conclusions derive from the analysis in this paper. First, weight, height, and BMI are highly persistent starting in early infancy. Second, majority of this persistence is attributable to time-invariant characteristics of children. This finding is crucial as it suggests that only interventions that alter these time-invariant attributes will have substantive, long-run effects on a child's anthropometric status. Thus, current policy interventions may, at best, have a marginal impact in the short run and, at worst, be destined to fail [see, e.g., Davis and Gebremariam (2010)]. Moreover, while it is difficult to say what these critical, time-invariant attributes are given the data at hand, we find some evidence that fetal nutrition—with mother's pre-pregnancy weight and weight gain during pregnancy, gestation age, birth status (singleton, twin, or higher-order birth), and birthweight as proxies—impacts the evolution of anthropometric measures over the early life cycle. However, time-invariant, *unobserved* attributes play a much more prominent role.

The remainder of the paper is organized as follows: Sect. 2 presents the methodology. Section 3 presents the ECLS-K sample analysis. Section 4 presents the ECLS-B sample analysis. Section 5 discusses the results of the analyses, and Sect. 6 concludes.

## 2 Methodology

We assess the extent and origins of persistence using a dynamic regression framework. This approach allows for the decomposition of persistence into various components reflecting state dependence, observed heterogeneity, and unobserved heterogeneity.

The simplest estimating equation is

$$y_{it} = \gamma y_{it-1} + \varepsilon_{it}, \qquad i = 1, \ldots, N; \; t = 1, \ldots, T, \tag{1}$$

where $y_{it}$ is a measure of anthropometric status for child $i$ at time $t$, $\varepsilon_{it}$ is a mean zero error term, and $T$ must be at least two (given observability of the initial observation, $y_{i0}$). Here, $\gamma$ reflects the overall level of persistence as it captures the entire association between past and current anthropometric status.

To decompose the overall persistence, we next incorporate *observed* heterogeneity into the model as follows:

$$y_{it} = \gamma y_{it-1} + x_{it}\beta + w_i\delta + \varepsilon_{it}, \qquad i = 1, \ldots, N; \; t = 1, \ldots, T, \tag{2}$$

where $x_{it}$ is a vector of observed, time-varying attributes of child $i$ at time $t$ and $w_i$ is a vector of observed, time-invariant attributes of child $i$. The change in the estimate of $\gamma$ from (1) to ( 2) reflects the portion of persistence attributable to observed heterogeneity.

Finally, we include observed time-varying heterogeneity and all sources (observed and unobserved) of time-invariant heterogeneity into the model as follows

$$y_{it} = \gamma y_{it-1} + x_{it}\beta + \alpha_i + \varepsilon_{it}, \qquad i = 1, \ldots, N; \; t = 1, \ldots, T, \tag{3}$$

where $\alpha_i$ is a child-specific fixed effect. In (3) $\gamma$ reflects the degree of *state dependence* as it captures the causal effect of past weight status on current weight status. The child-specific fixed effect, $\alpha$, reflects persistence in child anthropometric measures due to persistent observed and unobserved, child-specific heterogeneity (such as time-invariant environmental and genetic factors). In such models, $\beta$ represents the contemporaneous effects of the observed, time-varying regressors, whereas $\beta/(1-\gamma)$ represents the long-run effects of a permanent unit change in these variables.

Estimation of (3) is straightforward (assuming the model is correctly specified). Following Anderson and Hsiao (1981), (3) is first-differenced to eliminate $\alpha_i$. The first-differenced model is then estimated via instrumental variables since the first-differenced lagged dependent variable is necessarily correlated with the first-differenced error term. However, $y_{it-2}$ represents a valid instrument if $\varepsilon$ is serially uncorrelated.[3] The models are estimated by generalized method of moments (GMM).

[3] Although we proceed with $y_{it-2}$ as an instrument for the first-differenced lagged dependent variable, we note that this approach is limited since we reject the null hypothesis of no serial correlation for our sample. However, we also perform an exercise to assess the robustness of our estimates to omitted variable bias following Oster (2014) that is summarized in Sect. 3.2.

In addition to simply examining the coefficient estimates, we follow the logic in Ulrick (2008) and simulate probabilities (given estimates of the regression model) as follows:

$$\Pr(y_{iT} \geq y^* | y_{i0} \geq y_0). \tag{4}$$

Here, (4) represents the probability of a child having an anthropometric measure above $y^*$ in the terminal period conditional on an initial measure greater than or equal to some value $y_0$. For example, one might be interested in the probability of a child having a BMI above the 85th percentile in period $T$ conditional on being above the 85th percentile in the initial period, $t = 0$. These probabilities incorporate not just the coefficient directly related to persistence, $\gamma$, but also reflect persistence due to persistence in observed and unobserved determinants of child weight. Moreover, we can simulate counterfactual probabilities by altering the attributes of individuals.

Before detailing the simulations undertaken, note that upon estimating (3), the estimates of the child-specific fixed effects, $\alpha_i$, are given by

$$\widehat{\alpha}_i = \frac{1}{T} \sum_{t=1}^{T} \left[ y_{it} - \widehat{\gamma} y_{it-1} - x_{it} \widehat{\beta} \right], \qquad i = 1, \ldots, N. \tag{5}$$

These child-specific fixed effects estimates can then be decomposed into observed and unobserved time-invariant factors by estimating the following model using ordinary least squares:

$$\widehat{\alpha}_i = w_i \delta + \eta_i, \tag{6}$$

where $w_i$ now includes an intercept and $\eta_i$ is a mean zero error term.[4] Finally, given estimates of $\gamma$, $\beta$, and $\alpha$, we can obtain estimates of the idiosyncratic errors, $\varepsilon$ and $\eta$, using (3) and (6).

We now turn to the simulation of probabilities such as those given by (4) under the following counterfactual scenarios:

1. Own $y_{i0}$, own $x_{it}$, set $\varepsilon_{it} = 0$, and
   (a) replace $\alpha_i = \widehat{\overline{\alpha}}$, or
   (b) draw $\alpha_i \sim F(\alpha)$ where $F(\cdot)$ is the empirical distribution of $\alpha_i$.
2. Own $y_{i0}$, own $x_{it}$, set $\eta_i = 0$, set $\varepsilon_{it} = 0$, and
   (a) own $w_i$, or
   (b) draw $w_i \sim F(w)$ where $F(\cdot)$ is the empirical distribution of $w_i$.
3. Own $y_{i0}$, own $x_{it}$, own $w_i$, set $\varepsilon_{it} = 0$, and draw $\eta_i \sim F(\eta)$ where $F(\cdot)$ is the empirical distribution of $\eta_i$.
4. Own $y_{i0}$, own $\alpha_i$, set $\varepsilon_{it} = 0$, and
   (a) replace $x_{it} = \overline{x}_t$, or
   (b) draw $x_{i\cdot} \sim F(x_1, \ldots, x_T)$ where $F(\cdot)$ is the empirical joint distribution of $x_1, \ldots, x_T$.
5. Own $y_{i0}$, own $x_{it}$, own $\alpha_i$, and draw $\varepsilon_{i\cdot} \sim F(\varepsilon_1, \ldots, \varepsilon_T)$ where $F(\cdot)$ is the empirical distribution of $\varepsilon_{i\cdot}$..

[4] See, e.g., Hornstein and Greene (2012) and references therein for examples of two-step estimation approaches using a vector of estimated coefficients as a dependent variable.

6. Own $y_{i0}$, own $\alpha_i$, and draw $x_{i\cdot}, \varepsilon_{i\cdot} \sim F(x_1, \ldots, x_T, \varepsilon_1, \ldots, \varepsilon_T)$ where $F(\cdot)$ is the empirical joint distribution of $x_1, \ldots, x_T$, and $\varepsilon_1, \ldots, \varepsilon_T$.
7. Own $x_{it}$, own $w_i$, own $\alpha_i$, own $\varepsilon_{it}$, and
   (a) replace $y_{i0} = \overline{y}_{i0}$, or
   (b) draw $y_{i0} \sim F(y_{i0})$ where $F(\cdot)$ is the empirical distribution of $y_{i0}$.

Probabilities are obtained using 500 simulations. See "Appendix A" for further details.

Case 1 eliminates time-varying, unobserved heterogeneity, $\varepsilon_{it}$ , and assesses the impact of altering the distribution of time-invariant heterogeneity, $\alpha_i$. Case 1a eliminates all time-invariant heterogeneity by setting it to the sample mean of $\hat{\alpha_i}$ . Case 1b replaces actual time-invariant heterogeneity with a random draw from the empirical distribution of $\alpha_i$. Case 1b entirely breaks any correlation between the initial condition, $y_{i0}$, $x_{it}$, and time-invariant heterogeneity, $\alpha_i$. Put together, these scenarios highlight the relative importance of time-invariant heterogeneity in the origin of persistence in child anthropometric measures.

Case 2 continues to eliminate time-varying, unobserved heterogeneity, $\varepsilon_{it}$. However, time-invariant, unobserved heterogeneity, $\eta_i$, is now also eliminated; the observed component of time-invariant heterogeneity is then altered. Case 2a uses each observation's own time-invariant heterogeneity, $w_i$. Case 2b draws $w_i$ from the empirical distribution. Case 3 is similar to Case 2, but has individuals retain their time-invariant, observed heterogeneity, $w_i$, and alters the distribution of time-invariant, unobserved heterogeneity, $\eta_i$. Case 3a draws $\eta_i$ from the population empirical distribution. Altogether, Cases 2 and 3 permit the assessment of the relative importance of the observed and unobserved components of time-invariant heterogeneity in the persistence of child anthropometric outcomes.

Case 4 continues to eliminate time-varying, unobserved heterogeneity, $\varepsilon_{it}$, and assesses the impact of altering the distribution of time-varying, observed heterogeneity, $x_{it}$. Case 4a eliminates all time-varying heterogeneity. Case 4b replaces actual time-varying, observed heterogeneity with a random draw from the empirical distribution and succeeds in entirely breaking any correlation between the initial condition, $y_{i0}, \alpha_i$, and time-varying, observed heterogeneity, $x_{it}$. The scenarios in Case 4 complement the simulations performed in Case 1 as they speak to the relative importance of time-varying, observed heterogeneity in the persistence of child anthropometric outcomes.

Case 5 has individuals retain their time-varying, observed attributes, $x_{it}$, time-invariant attributes, $\alpha_i$ and $y_{i0}$, but alters the distribution of time-varying, unobserved heterogeneity, $\varepsilon_{it}$ by drawing $\varepsilon_{i\cdot}$ from the empirical distribution. Case 6 has individuals retain only their time-invariant attributes, $\alpha_i$ and $y_{i0}$. All time-varying heterogeneity is sampled. In particular, $x_{i\cdot}$ and $\varepsilon_{i\cdot}$ are drawn from the population empirical distribution. Thus, Cases 5 and 6 address the relative importance of the observed and unobserved components of time-varying heterogeneity in the persistence of childhood anthropometric outcomes.

Finally, Case 7 has individuals retain their time-varying, observed attributes, $x_{it}$, time-invariant, observed heterogeneity, $w_i$, and time-invariant, unobserved heterogeneity, $\alpha_i$, and assesses the importance of altering the initial condition, $y_{i0}$. While Case 7a replaces the initial condition with the sample mean of $y_{i0}$, Case 7b replaces the actual initial condition with a random draw from the empirical distribution of $y_{i0}$.

## 3 ECLS-K sample analysis

### 3.1 Data

We use data from the restricted version of the ECLS-K. Collected by the U.S. Department of Education, the ECLS-K surveys a nationally representative cohort of children throughout the USA in fall and spring kindergarten, fall and spring first grade, spring third grade, spring fifth grade, and spring eighth grade. The sample includes data on over 20,000 students who entered kindergarten in one of roughly 1000 schools during the 1998–1999 school year. In addition to family background information, height and weight measures are available for children in each round, as well as information on birth weight.

Our final sample consists of children for whom we have valid measures of age, gender, height, and weight.[5] From the information on height and weight of the children, we obtain $z$-scores for weight, height, and BMI. Note that $z$-scores and percentiles are based on CDC 2000 growth charts; these are age and gender specific and are adjusted for normal growth, and percentiles are based on the underlying reference population.[6] The estimation utilizes data from five waves: fall kindergarten, spring first grade, spring third grade, spring fifth grade, and spring eighth grade.[7] The sample is a balanced a panel of roughly 9160 children.[8]

The following time-invariant covariates are included in the regression model: gender, race/ethnicity (white, black, Hispanic, Asian, and other), birthweight, indicator for premature birth, indicator for being born in the USA, indicator for being a native English speaker, city type (urban, suburban, or rural), region (northeast, midwest, south, and west), mother's education (less than high school, high school/GED, some college, 4-year college degree, and more than 4 years of college), mother's age at first birth, mother's marital status at birth, mother's height, mother's pre-pregnancy weight, mother's weight gain during pregnancy, gestation age, birth status (singleton, twin, or higher-order birth), indicator for attending non-parental prekindergarten, indicator for mother's labor force participation during infancy, indicator for mother's participation in WIC (Women, Infants, and Children) during pregnancy, indicator for mother's participation in WIC during infancy, indicator for mother's participation in temporary assistance for needy families (TANF) during infancy, indicator for participation in food stamp program (FSP) during infancy, and indicator for attending full-day kindergarten.[9]

[5] The initial sample size of the ECLS-K is 21,260. After cleaning age, weight, and height as described in Millimet and Tchernis (2013b, Appendix C), and due to sample attrition, the sample size falls to 9360 in the final wave of the data. Restricting the sample to a balanced panel reduces the sample size to approximately 9160. This is the final sample size per wave in the analysis.

[6] $z$-scores and their percentiles are obtained using the *-zanthro-* command in Stata.

[7] The survey design is troublesome in that the ECLS-K contains irregularly spaced waves. To minimize the issue, we omit the spring kindergarten wave and thus each period conceptually represents roughly a 2-year window.

[8] Note, all sample sizes are rounded to the nearest 10 per NCES restricted data regulations.

[9] FSP was renamed the supplemental nutrition assistance program (SNAP) in October 2008. Since the data predate this change, we refer to the program as FSP.

The following time-varying covariates are included in the regression model: an index of socioeconomic status (SES), indicator for the household being in poverty, the number of children's books in the household, household size, family type (two parents plus siblings, two parents and no siblings, one parent and siblings, one parent and no siblings, and other), mother's labor force participation status (full time, part time, and not working), mother's weight, indicator for mother absent from the household, indicator of current TANF participation, indicator of current FSP participation, indicator for health insurance, hours spent watching television during the school week, hours spent watching television during the weekend, indicator for household rules regarding television watching, days per week household eats breakfast together, days per week household eats dinner together, indicator for household food security (household never worried about running out of food), neighborhood safety (very safe, somewhat safe, and not safe), and percent of minority students in class at school. For all covariates (except gender, age, height, and weight), we include dummy variables for missing observations.

Table 1 presents descriptive statistics of the above time-varying and time-invariant covariates. Although we use data from five waves of the ECLS-K, Table 1 presents means and standard deviations (in parentheses) for the final sample after restrictions for the first wave (see first column of Table 1) and the initial sample before restrictions (see the second column of Table 1). Although we do not discuss the summary statistics in detail for brevity, the sample restrictions we implement, as well as attrition, do not lead to a drastically different sample, at least at the mean, except a few cases such as mother's marital status at birth, the number of children's books in household, indicator for child's health insurance status, indicator for neighborhood being very safe, and the indicator for the household never been worried about running out of food.

### 3.2 Results

Table 3 displays the results from estimation of Eqs. (1), (2), and (3) for weight, height, and BMI $z$-scores, respectively. In addition to reporting estimates of the coefficient on the lagged outcome, $\gamma$, we report the first-stage Kleibergen and Paap (2006) Wald rk $F$-statistic, the Kleibergen and Paap (2006) rk test of underidentification, and a test of endogeneity. The first two tests are designed to detect any issues associated with weak instruments. Finally, recall that the estimate of $\gamma$ from (1) reflects the overall level of persistence, the change in the estimate of $\gamma$ moving from (1) to (2) captures the portion of persistence explained by the observable covariates, and the change moving from (2) to (3) reflects the portion of persistence explained by unobserved, time-invariant factors.[10]

Table 4 presents the dynamic simulations based on (4), which affords further analysis of the sources of persistence, especially in the upper tail of the distribution of

[10] Also, we evaluate the robustness of our estimates in Eq. (3) to omitted variables bias following Oster (2014). Although the estimates of the implied ratio of the selection on unobservables relative to selection on observables, $\delta$, are less than 1, all but one of estimates of $\delta$ are negative. This suggests that, although a relatively small amount of selection on unobservables (relative to observables) is needed to yield a treatment effect of zero (i.e., $\gamma = 0$), this must operate in the opposite direction to be the case. The results of the Oster (2014) approach are available upon request from the authors.

anthropometric measures, and the role of time-varying and time-invariant observed attributes. As noted earlier, the simulations are based on the estimates of the fixed effects specification given in (3), along with the subsequent estimates of the fixed effects and their decomposition given in (5) and (6). From the regression results, for each outcome, we simulate three sets of probabilities:

1. $\Pr(y_{iT} \geq$ 85th percentile $|\, y_{i0} \geq$ 85th percentile),
2. $\Pr(y_{iT} \geq$ 95th percentile $|\, y_{i0} \geq$ 95th percentile), and
3. $\Pr(y_{iT} \geq$ 85th percentile $|\, y_{i0} \leq$ 50th percentile),

where period $T$ denotes spring eighth grade and period 0 corresponds to fall kindergarten. Note the percentile outcomes are based on the underlying reference population used in the CDC 2000 growth charts, *not* the current sample. Thus, the 85th and 95th percentiles correspond to usual cutoffs for overweight and obese when examining BMI. Finally, Table 4 also presents the *benchmark* probability, which is the empirical probability observed in the data (i.e., the sample probability as opposed to an estimate), for comparison.

### *3.2.1 Weight*

Panel I of Table 3 displays the regression results for weight $z$-scores. The estimates of $\gamma$ across the three specifications are 0.931, 0.932, and 0.775 (standard errors are 0.003, 0.003, and 0.067, respectively) in Columns (1), (2), and (3), respectively. Each coefficient estimate is statistically significant at the 1% confidence level, and all three specifications are strongly identified. The estimates of $\gamma$ above 0.9 indicate a substantial degree of childhood anthropometric persistence. Moreover, the fact that the estimate of $\gamma$ does not change moving from Column (1) to Column (2) implies that our lengthy vector of time-varying and time-invariant observed factors explains *none* of the persistence in weight status for primary-school-aged children. Thus, while we find extreme persistence from one period to the next, this persistence is not attributable to or explained by characteristics typically observed by policy-makers or health practitioners.

Moving to the specification in Column (3), which replaces the time-invariant observed factors with child-level fixed effects and thereby accounts for all time-invariant attributes of the child, the estimate of $\gamma$ falls to 0.775, a decline of roughly 17% from 0.93 in Columns (1) and (2). This implies that time-invariant, unobserved factors explain about 17% of the observed persistence in weight $z$-scores. Examples of such factors include genetic endowments, prior health shocks determined *in utero* or during infancy, time-invariant environmental factors such as the presence of grocery stores or outdoor amenities, etc.

In sum, we find that overall persistence is fairly extreme as a one standard deviation increase in weight is associated with roughly nine-tenth of a standard deviation increase in the subsequent period. However, time-varying and time-invariant observed attributes explain none of this persistence. Furthermore, time-invariant unobserved factors also explain very little of the persistence (less than one-fifth). Thus, much of the persistence in child weight is attributable to state dependence, which implies that early interventions that are successful in reducing child weight will have long-run effects.

Unfortunately, since our covariates explain little of the variation in weight, identifying such early interventions may be difficult.[11]

Table 4 displays the simulation results for $\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \geq$ 85th percentile), $\Pr(y_{iT} \geq$ 95th percentile $|y_{i0} \geq$ 95th percentile), and $\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \leq$ 50th percentile) for weight, height, and BMI $z$-scores in Columns (1)–(3), (4)–(6), and (7)–(8), respectively.

*Simulated probabilities:* $Pr(y_{iT} \geq$ 85th *percentile* $|y_{i0} \geq$ 85th *percentile)*

Focusing on Column (1) of Table 4, the benchmark probability is 0.84. In other words, 84% of children above the 85th percentile in the initial period remain above the 85th percentile in the terminal period in our sample. This is consistent with a high degree of persistence in weight in the upper tail of the distribution. We turn to the simulations to explore the sources of this persistence.

Panel I in Table 4 contains the simulated probabilities when time-varying unobservables are ignored (i.e., $\varepsilon_{it} = 0$ for all $i, t$) and time-invariant heterogeneity is altered first by removing it entirely (i.e., setting $\alpha$ to the sample mean of $\widehat{\alpha}$) and then by retaining the heterogeneity in $\alpha$, but breaking its correlation with $x$ and $y_0$ by giving each child a random draw from the empirical distribution of $\widehat{\alpha}$. In the first case, the conditional probability of staying above the 85th percentile falls to about 0.753, while it falls to about 0.576 in the second case. The fact that the conditional staying probability drops noticeably from the benchmark in the second case, but only marginally in the first case, indicates that the *correlation* between $\alpha$ and the time-varying covariates exerts a greater influence on persistence than the *variation* in $\alpha$ across children. In fact, the correlation between time-invariant heterogeneity and the time-varying covariates explain a little over 30% of total persistence in weight status (i.e., $1 - (0.576/0.84) = 0.314$). Moreover, since the results in Table 1 indicate that the time-varying, observed covariates, $x$, have little explanatory power, it is the correlation between $\alpha$ and the initial condition, $y_0$, that explains nearly one-third of the total persistence in weight status. In other words, children with high initial conditions—measured by weight $z$-scores upon kindergarten entry—also have high values of $\alpha$, and this combination is responsible for one-third of the conditional staying probability over the span of kindergarten through eighth grade.

Panels II and III in Table 4 assess whether the importance of $\alpha$ is driven by time-invariant observed factors, $w$, or unobserved factors, $\eta$. The first simulation in Panel II sets $\eta$ equal to zero and keeps $w$ at its actual value. The result is very similar to the first case in Panel I, where $\alpha$ is set equal to its sample mean. In this case, the conditional staying probability is 0.727, implying that the results in Panel I were driven by having set $\eta$ to its sample mean. When children were instead given a random draw for $w$ from its empirical distribution, the probability changes only modestly to 0.703. Again, this is consistent with the results in Table 3 where we found little explanatory

[11] The full set of results are available upon request. While some estimated coefficients are statistically significant at conventional levels, the magnitudes are quite small; even the long-run effects of permanent changes in the covariates, given by $\beta/(1 - \gamma)$, are quite small. That said, while our covariate set does include a wide array of the usual family background variables, we do not have information on many recent interventions designed to combat obesity, such as education efforts, healthy food programs, and efforts to promote physical activity. We also do not have data on parents' height or weight. We return to the issue of parental anthropometric status later.

power for the time-invariant, observed covariates. In Panel III, however, when children retain their own observed factors, $x$ and $w$, but receive a random draw for $\eta$ from its empirical distribution, the conditional staying probability falls to 0.591. As such, it is the correlation between time-invariant, *unobserved* factors and the initial condition, $y_0$, that is responsible for roughly one-third of the conditional staying probability. In other words, children with high initial conditions also have high values of $\eta$, and this combination is responsible for one-third of the persistence in weight from kindergarten through eighth grade.

Panels IV, V, and VI in Table 4 report the simulated probabilities obtained when children retain their time-invariant heterogeneity, $\alpha$, but receive draws of either time-varying, observed covariates, $x$, or unobserved attributes, $\varepsilon$, or both from their respective empirical distributions. The results indicate no impact from altering either, which is consistent with the prior results in Table 3.

Lastly, Panel VII in Table 4 presents the simulated probabilities when the initial condition, $y_0$, is altered by first, setting it equal to the sample mean for all $i, t$ and second, by giving each child a random draw from the empirical distribution of $y_0$. In the first case, by removing all heterogeneity in the initial condition (and thus, entirely breaking its correlation with $\alpha$ and $x$), the conditional probability of staying above the 85th percentile falls from 0.840 to about 0.337. When children are instead given a random draw from the empirical distribution of $y_0$, the conditional staying probability falls only to 0.583.

In sum, the simulations indicate that about one-third of the conditional probability of staying above the 85th percentile for weight is due to *time-invariant*, *unobserved* risk factors such as genetic endowments, early-life health shocks, time-invariant environmental factors. In spite of the increased role of time-invariant, unobserved heterogeneity, a greater portion of the persistence in weight status is still attributable to state dependence as evidenced by the regressions in Table 1. The fact that about two-thirds of persistence in weight status is due to state dependence is encouraging in that early interventions, to the extent that they are successful in reducing weight prior to kindergarten, can have long-run effects on weight during middle school.

*Simulated probabilities:* $Pr(y_{iT} \geq$ 95th *percentile* $|y_{i0} \geq$ 95th *percentile)*

Column (2) of Table 4 displays the analogous results for $\Pr(y_{iT} \geq$ 95th percentile $|y_{i0} \geq$ 95th percentile). Compared to the results in Column (1) of Table 4, three primary differences emerge. First, the benchmark probability of 0.762 is lower for the full sample. Second, the vast majority of the persistence is due to variations in time-invariant heterogeneity, $\alpha$, across children even more so than in Column (1). In fact, the conditional probability of staying above the 95th percentile falls to about 0.134 when the time-invariant heterogeneity is altered by removing it entirely (i.e., setting $\alpha$ to the sample mean of $\widehat{\alpha}$) in Panel I of Column (2). State dependence, as well as time-varying factors, $x$ and $\varepsilon$, do not play much of a role in explaining persistence in the extreme upper tail. Third, unlike in Column (1) of Table 4, we find that setting $\eta$ to zero in Panel II results in a relatively greater reduction in the conditional staying probability than in Panel III when $\eta$ is replaced by random draws from different empirical distributions (i.e., about 45 vs. about 30%). This indicates that giving children initially above the 95th percentile an average draw from the distribution of $\eta$ (i.e., setting $\eta$ to zero) is

sufficient to bump most of these children below the 95th percentile by the terminal period, whereas this is not sufficient when using the 85th percentile as the threshold. *Simulated probabilities:* $Pr(y_{iT} \geq 85\text{th } percentile \mid y_{i0} \leq 50\text{th } percentile)$

Finally, Column (3) of Table 4 presents the results for $\Pr(y_{iT} \geq 85\text{th percentile} \mid y_{i0} \leq 50\text{th percentile})$. This case illuminates factors associated with relatively extreme weight gain during early childhood (i.e., sizable upward mobility as opposed to persistence). In terms of the benchmark case, the probability of moving from below the median at kindergarten entry to above the 85th percentile by the end of eighth grade is roughly 12%.

Turning to the simulations, we obtain a few noteworthy findings. First, time-varying factors, $x$ and $\varepsilon$, continue not to play any meaningful role. Second, replacing $\alpha$ with the sample mean reduces the probability of crossing the 85th percentile conditional on starting below the median to zero. Replacing $\alpha$ with a random draw from different empirical distributions roughly doubles the probability of crossing the 85th percentile relative to the benchmark. Similarly, replacing the initial condition with a random draw from the empirical distribution of $y_0$ roughly doubles the conditional probability of crossing the 85th percentile relative to the benchmark. Put together, these results imply that children initially below the median tend to have favorable values of $\alpha$. Specifically, $\alpha$ is not randomly distributed in the population, but rather has a positive (partial) correlation with the initial condition, $y_0$. Only the few children with extremely unfavorable draws of $\alpha$, despite being below the median in the initial period, experience extreme upward mobility. Moreover, if $\alpha$ were randomly assigned, the probability of moving from below the median to above the 85th percentile would roughly double. This is a testament to the importance of time-invariant factors (not state dependence), in determining the probability of extreme upward mobility in weight status.

Third, the effect of randomly assigning $\alpha$ is due to randomly assigning time-invariant, unobserved factors, $\eta$. Randomly assigning the time-nvariant, observed factors, $w$, has little impact on the probability of extreme upward mobility. Moreover, removing time-nvariant, unobserved factors by setting $\eta$ to zero reduces the probability of extreme upward mobility to nearly zero. The implication is that children below the median tend to have favorable draws of $\alpha$, which really means favorable draws of time-nvariant, unobserved factors, $\eta$.

Overall, the evidence presented thus far points to an important role for state dependence in explaining persistence in weight status on average but less so in the upper right tail of the weight distribution, where time-nvariant, unobserved heterogeneity assumes a greater role. Next we turn to the analysis of height.

### *3.2.2 Height*

While height *per se* is not a policy concern in the USA, it is interesting to compare the dynamics of height with those of weight. In addition, it is useful to examine the individual components of BMI prior to assessing BMI $z$-scores in the next section.

Panel II of Table 1 displays the results for height $z$-scores. The estimates of $\gamma$ across the first two specifications are very similar to those using weight $z$-scores, namely 0.937 and 0.936. (Standard errors are 0.004 and 0.004, respectively.) However, the estimate of $\gamma$ falls to 0.603 (standard error is 0.048) in the fixed effect specification

(compared to 0.775 in Table 1). Similar to Panel I of Table 1, the estimate of $\gamma$ is statistically significant at the $p < 0.01$ confidence level, and all three specifications are strongly identified. The estimate of $\gamma$ barely changes when we include time-arying and time-nvariant observed attributes, and the estimates of $\gamma$ above 0.9 in the first two specifications indicate a substantial degree of persistence in height. Thus, as for weight in Panel I of Table 1, while height is quite persistent from one period to the next, this is not attributable to or explained by observed characteristics.

However, in contrast to weight $z$-scores, the inclusion of child-level fixed effects explains about 36% (i.e., $1 - (0.603/0.936) = 0.356$) of the overall persistence in child height (versus only 17% for weight $z$-scores). This is perhaps not surprising as unobserved biological factors—most noticeably, parental height—are not included in our set of observed covariates. The fact that time-nvariant, unobserved attributes account for a greater share of the persistence in height implies that state dependence, and thus the long-run impact of successful, early interventions—that do not alter relevant, time-nvariant, unobserved attributes—is diminished.

Columns (4)–(6) of Table 4 present the analogous set of simulation results for height $z$-scores. In terms of the benchmark probabilities, a few differences emerge relative to the prior results for weight. First, the benchmark probabilities are lower for height than the corresponding probabilities for weight. For example, $\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \geq$ 85th percentile) and $\Pr(y_{iT} \geq$ 95th percentile $|y_{i0} \geq$ 95th percentile) are 0.606 and 0.467, respectively, for height; these probabilities are 0.840 and 0.762, respectively, for weight. Thus, persistence in the upper half of the distribution is lower, albeit still high, for height. Second, while there may exist more mobility in terms of height, extreme upward mobility for height is less common than for weight—$\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \leq$ 50th percentile) is 0.030 for height and 0.118 for weight.

Turning to the simulations, a few patterns emerge. First, while the time-varying factors, $x$ and $\varepsilon$, have a bit more impact on height than weight, their combined effect is still modest. In Columns (4)–(6) of Table 4, replacing $x$ and/or $\varepsilon$ with different values increases the conditional staying probabilities in all cases (See Panels IV through VI). This indicates that, on average, children initially above the median tend to have less favorable (in terms of raising height) time-varying attributes, partially offsetting the child's height in the initial period.

Second, most of the persistence in height is attributable to time-invariant factors captured by $\alpha$, but with different patterns when compared to prior results for weight. For instance, in Columns (4)–(6) of Table 4, we find that replacing $\alpha$ with the sample mean drops the conditional staying probabilities above the 85th and 95th percentiles to zero. Further analysis reveals that this stems from the unobserved component captured by $\eta$; varying the time-invariant, observed component, $w$, has little effect. This implies that children in the upper tail of the height distribution upon entry to kindergarten possess time-invariant, unobserved attributes that tend to keep them in the upper tail. Replacing these attributes with the sample mean or a random draw of $\alpha$ essentially guarantees these children will fall out of the upper tail by the end of eighth grade. Replacing the unobserved component of the fixed effects, $\eta$, with a random draw similarly reduces the conditional staying probabilities, but not as much; the probabilities fall to around 0.256 and 0.108 in Columns (4) and (5) of Table 4, respectively. In addition, replacing the initial condition with a random draw from the empirical distribution

reduces the conditional probabilities of staying above the 85th and 95th percentiles to 0.241 and 0.435. This is perhaps not surprising as genetics and early biological factors presumably play a large role in determining child height.

Third, Column (6) of Table 4 suggests that extreme upward mobility in height is rare since children initially below the median have unfavorable draws of time-invariant, unobserved heterogeneity, $\eta$. Replacing $\eta$ with its sample average would eliminate extreme upward mobility entirely as the few cases of observed extreme upward mobility are due to a handful of children having very favorable values of $\eta$ despite being below the median upon entry into kindergarten. On the other hand, replacing $\eta$ with a random draw would increase extreme upward mobility almost fourfold. This pattern is similar to the impact of replacing the initial condition with a random draw where the conditional probability increases from 0.030 to 0.153. In sum, although on average, state dependence is still important in explaining persistence in height (i.e., almost two-thirds), persistence in the upper tail of the height distribution is almost entirely attributable to time-invariant, unobserved heterogeneity. We now turn to the analysis of BMI.

### *3.2.3 BMI*

Panel III of Table 3 presents the regression results for BMI $z$-scores. For the full sample, the estimates of $\gamma$ across the first two specifications are very similar to those in Panels I and II of Table 3, namely 0.912 and 0.911. (Standard errors are 0.004 and 0.005, respectively.) However, the estimate of $\gamma$ now falls to 0.217 (standard error is 0.015) in the fixed effect specification (compared to 0.775 and 0.603 in Panels I and II of Table 3, respectively). As in Panels I and II of Table 3, the estimate of $\gamma$ is statistically significant at the $p < 0.01$ confidence level and all three specifications are strongly identified. Also, the estimate of $\gamma$ barely changes when we include time-varying and time-invariant observed attributes, and the estimates of $\gamma$ above 0.9 in the first two specifications indicate a substantial degree of persistence. Thus, as with weight and height $z$-scores, while persistence from one period to the next in BMI $z$-scores is high, it is not attributable to or explained by observed characteristics.

While the first two specifications differ little across Panels I, II, and III of Table 3, the results from the fixed effect specification have more pronounced differences. As noted above, time-invariant, unobserved factors account for roughly 17% of the total persistence in weight $z$-scores and 36% for height $z$-scores. In sharp contrast and perhaps, surprisingly, the fixed effects account for nearly 80% of total persistence in BMI. The economically and statistically meaningful drop in the estimate of $\gamma$ implies a substantially smaller role for state dependence in the persistence of child BMI. Consequently, the long-run impact of early interventions—that do not alter relevant, time-invariant, unobserved attributes—on BMI is quite small. For example, a *one-time* intervention that reduces a child's BMI prior to kindergarten entry by one standard deviation, *ceteris paribus*, is expected to have essentially no impact on BMI in spring eighth grade. A *permanent* intervention that reduces a child's BMI by 0.10 standard deviations *every period* will only result in a long-run decrease in the child's BMI of roughly 0.13 standard deviations. This has profound implications for the types of policies one should pursue if the objective is to reverse the obesity epidemic.

Columns (7)–(9) of Table 4 display the simulation results for BMI $z$-scores. In Columns (7) and (8), the benchmark probabilities lie in between the conditional staying probabilities for weight and height reported in the corresponding Columns (1)–(2) and Columns (4)–(5). Furthermore, the benchmark probabilities are consistent with the high degree of persistence in BMI documented earlier. For example, the conditional probability of staying above the 85th percentile is 0.746 (see Column (7) of Table 4) and 0.715 for staying.

Turning to the simulated probabilities, a few findings stand out. First, altering the values of the time-invariant components in Panels I, II, and III of Columns (7) and (8) in Table 4 yields results that are qualitatively similar to those reported in Columns (4) and (5) for height. In particular, in Panel I, we find that replacing $\alpha$ with the sample mean reduces the conditional probability of staying above the 85th and 95th percentiles to zero in nearly every case. Moreover, this is predominantly due to the salient role of time-invariant, unobserved factors, $\eta$. However, the impact of replacing the initial condition with the mean of $y_0$ as well as a random draw from the empirical distribution is similar to the reduction in the conditional probability that results from replacing $\alpha$ with a random draw. Variation in time-invariant, observed factors, $w$, explains a modest amount of variation in the conditional probability of staying above the 85th percentile [see Column (7)], but not when using the 95th percentile as the threshold [see Column (8)]. Thus, the results are consistent with children in the upper part of the BMI distribution possessing less favorable time-invariant factors, particularly those unobserved. Second, as in all the analysis of weight and height, we find very little role for variation in time-varying factors, either observed or unobserved.

Column (9) of Table 4 presents the results for $\Pr(y_{iT} \geq \text{85th percentile} \,|\, y_{i0} \leq \text{50th percentile})$. In terms of the benchmark probabilities for extreme upward mobility, we obtain higher probabilities for BMI than for either weight or height. For example, the probability of having a BMI above the 85th percentile in the terminal period conditional on entering kindergarten below the median is 0.142 [see Column (9)], while the corresponding figures are 0.118 and 0.003 for weight and height, respectively.

Turning to the simulations in Column (9) of Table 4, we obtain a few noteworthy findings. First, time-varying factors, $x$ and $\varepsilon$, continue to not play an important role. Second, replacing $\alpha$ with the sample mean reduces the probability of crossing the 85th percentile conditional on starting below the median to zero, just as in Columns (3) and (6) of Table 4. Replacing $\alpha$ with a random draw from different empirical distributions increases the probability of crossing the 85th percentile roughly two-and-a-half-fold relative to the benchmark (see Column (9) of Panel I in Table 4). Also, replacing the initial condition with a random draw from the empirical distribution yields a similar two-and-a-half-fold increase in the conditional probability of crossing the 85th percentile. Together, these results continue to imply that children initially below the median tend to have favorable values of $\alpha$. Only a few children with extremely unfavorable draws of $\alpha$, despite being initially below the median, experience extreme upward mobility. Moreover, if $\alpha$ were randomly assigned, the probability of moving from below the median to above the 85th percentile would increase substantially.

Third, the effect of altering $\alpha$ is due to altering the time-invariant, unobserved factors, $\eta$. However, as in Column (7), the time-invariant, observed factors, $w$, explain a modest amount of the variation in the probability of extreme upward mobility. Specifi-

cally, whereas removing time-invariant, unobserved factors by setting $\eta$ to zero reduces the probability of extreme upward mobility to nearly zero for weight and height, this is not the case for BMI as the probability varies from roughly 4–5%.

Summarizing, while persistence in BMI, on average, is largely attributable to time-invariant, unobserved heterogeneity, it is almost entirely driven by the same in the upper tail of the BMI distribution. State dependence plays the least role in explaining persistence in BMI compared to height and weight. An interesting insight that follows from the ECLS-K sample analysis is that examining persistence in BMI alone can potentially mask the importance of state dependence relative to time-invariant, unobserved heterogeneity as the former plays a bigger role in the persistence of the components of BMI (i.e., weight and height). The remainder of this paper undertakes a complementary analysis of the ECLS-B sample, comparing and contrasting its results to those of ECLS-K sample analysis.

## 4 ECLS-B sample analysis

### 4.1 Data

To explore the early-life origins of childhood anthropometric persistence, we use data from the restricted version of the ECLS-B. Collected by the U.S. Department of Education, the ECLS-B collects information on a nationally representative cohort of children born in 2001 at 9 months of age, 2, 4, and 5 years. As with the ECLS-K, our final sample consists of a balanced sample of children for whom we have valid measures of age, gender, height, and weight.[12] Given the age of the sample, we convert weight into $z$-scores; height is, however, measured in centimeters.

The following time-invariant covariates are included in the regression model: gender, race/ethnicity (white, black, Hispanic, Asian, and other), mother's age at first birth, birthweight indicators (normal or low), indicator for intrauterine growth retardation (less than 10, 10–24, 25–49, 50–75, 76–89, and 90% and above),[13] indicator for premature birth, indicator for birth status (singleton, twin, or higher-order birth), mother's height, mother's weight prior to pregnancy, mother's weight gain during pregnancy, indicator for prenatal care (inadequate, intermediate, adequate, or adequate plus), indicator for maternal prenatal vitamin consumption within the 3 months preceding conception, indicator for maternal prenatal vitamin consumption during the first trimester, indicator for maternal smoking within the 3 months preceding conception, indicator for maternal smoking within the third trimester, indicator for whether mother has smoked more than 100 cigarettes in her lifetime, indicator for maternal alcohol consumption within the 3 months preceding conception, the number of current smokers in the household, region (northeast, midwest, south, and west), city type

[12] The possible sample size is roughly 6950; the initial sample size in the first wave is about 10,700. After restricting the sample to those with valid data on age, gender, height, and weight, the sample size reduces to approximately 5,450. This is the final sample size per wave used in the regression analysis. Note that all sample sizes are rounded to the nearest 50 per NCES restricted data regulations for the ECLS-B.

[13] Intrauterine growth retardation measures the ratio of birthweight to predicted weight based on gestation age.

(urban, suburban, or rural), indicator for mother's participation in WIC during pregnancy, indicator for mother's participation in WIC during infancy, and scores on infant mental and motor assessments administered at 9 months.

The following time-varying covariates are also included: age, mother's education (less than high school, high school/GED, some college, 4-year college degree, and more than four years of college), an index of SES status, indicator for the household being in poverty, the number of children's books in the household, household size, family type (two parents plus siblings, two parents and no siblings, one parent and siblings, one parent and no siblings, and other), indicator for biological mother present, indicator for biological father present, indicator for no father present, indicator for no mother present, indicator for parental respondent's marital status, indicator of current TANF participation, indicator of current FSP participation, indicator for health insurance, indicator for current medicaid participation, indicator for current WIC participation, indicator for household food security (household never worried about running out of food), hours per day spent watching television during the school week, indicator for household rules regarding television watching, neighborhood safety (very safe, somewhat safe, and not safe), mother's labor force status (full time, part time, and not working), indicators for primary child care arrangement (parents, other relatives, non-relatives, center-based care, or Head Start), indicator for school enrollment, indicator if English is the primary home language, and mother's weight. For all covariates (except gender, age, height, and weight), we include dummy variables for missing observations.

Similarly, Table 2 presents descriptive statistics of the time-varying and time-invariant covariates in the ECLS-B sample. Although we use data from four waves of the ECLS-B, Table 2 presents means and standard deviations (in parentheses) for the final sample after restrictions for the first wave (see first column of Table 2) and the initial sample before restrictions (see the second column of Table 2). Again, at least at the mean, the summary statistics do not suggest that the sample restrictions we implement result in a significantly different sample.

Also, by comparing the first columns of Tables 1 and 2, there are no significant observed differences between the ECLS-K and ECLS-B samples' covariates at the mean, except for a few attributes such as the percent living in urban, suburban, and rural areas, the percent of households with two parents and siblings, the percent of households with two parents and no siblings, the percent of mothers that do not work, and the percent of households that never worried about running out of food.

### 4.2 Results

The results for the ECLS-B sample are presented in Tables 5 and 6. Table 5 displays the regression estimates, while Table 6 presents the simulation results.

In terms of the coefficient estimates, the results in Panel I of Table 5 using weight $z$-scores are fairly similar to those obtained using the ECLS-K when child-specific fixed effects are omitted. Specifically, the estimates of $\gamma$ are statistically significant and range from 0.870 to 0.873. The fact that the estimate of $\gamma$ does not change moving from (1) to (2) implies that our lengthy vector of time-varying and time-invariant

observed factors explains none of the persistence in weight status for infants and young children. Given the additional time-invariant controls available in the ECLS-B, this result is striking. Moreover, the estimates of $\gamma$ near 0.9 indicate a substantial degree of persistence even prior to kindergarten. However, unlike in the ECLS-K sample, the inclusion of child-level fixed effects explains the majority of this persistence. Here, the estimate of $\gamma$ falls to 0.124 (standard error of 0.013). This implies that time-invariant, unobserved factors explain about 85% of the observed persistence in weight $z$-scores during early childhood. Recall that, in contrast, only 17% of observed persistence in weight $z$-scores during primary school (ECLS-K sample) is due to time-invariant, unobserved heterogeneity. Again, given that we observe many more time-invariant attributes of children in the ECLS-B, this is a startling result.

Panel II of Table 5 displays the corresponding regression results for height. Four interesting patterns emerge. First, persistence in height in the models not controlling for any other covariates—based on the specification in (1)—is of a much smaller magnitude than found in the ECLS-K when assessing height for older children or in the ECLS-B when assessing weight. Second, when controlling for observed heterogeneity—based on the specification in (2)—persistence actually increases by about 15%. This is consistent with a negative correlation between the initial condition for height, $y_0$, which is really "length" at 9 months of age, and observed heterogeneity associated with greater height. Finally, when child-level fixed effects are included, the estimates of $\gamma$ become *negative* and statistically insignificant. Thus, *all* of the persistence in child height up to age five is attributable to *time-invariant*, *unobserved* heterogeneity.

Columns (1)–(3) and (4)–(6) of Table 6 report the results of the same simulations performed using the ECLS-K. In the interest of brevity, we only briefly summarize the results of the simulations. First, time-varying attributes, both observed and unobserved, have no effect on persistence. Given the lengthy vector of attributes, as well as the plethora of time-varying, unobserved attributes captured by $\varepsilon$, this continues to be a noteworthy finding.

Second, time-invariant heterogeneity continues to play a prominent role in understanding persistence in child weight and height. For weight, replacing $\alpha$ with its sample mean explains virtually all persistence through age five. Moreover, replacing the fixed effect of a child initially below the median with the sample mean roughly doubles the probability that the child's weight will exceed the 85th percentile at age five. However, replacing the fixed effect of a child initially below the median with the sample mean does not alter the probability that the child's height will exceed the 85th percentile at age five. In addition, for weight, replacing the initial condition with either the mean of $y_0$ or a random draw from the empirical distribution reduces the conditional probabilities of staying above the 85th and 95th percentiles with sizes comparable to the impact of altering the time-invariant, unobserved heterogeneity. Similar patterns in the reduction in conditional probabilities due to altering the initial condition are observed for height.

Third, time-invariant, *observed* attributes play a more prominent role, particularly for height, in explaining persistence up to age five than in the ECLS-K sample of primary school children. This could be attributed to two sources. On the one hand, the time-invariant, observed attributes may play a more important role in the determination of child weight and height prior to age five. On the other hand, the vector of controls is

not identical across the two data sources. Examining the results of (6) for the ECLS-B sample, the most important covariates relate to birthweight, birth status (i.e., singleton, twin, or higher-order birth), intrauterine growth retardation, breastfeeding duration, mother's height, and mother's weight gain during pregnancy. That said, as measured by the $R^2$, only 19% (22%) of the variation in $\widehat{\alpha}$ is explained by the covariates included in (6) when examining weight (height) for the ECLS-B sample and even smaller for the ECLS-K sample (i.e., 4.5, 5, and 6% for weight, height, and BMI, respectively).[14]

## 5 Discussion

This section compares and contrasts the analyses of the ECLS-K and ECLS-B samples and summarizes the salient findings of this paper.

First, the results from the sample of children aged five and younger in the ECLS-B are consistent with the sample of primary school children in the ECLS-K: Persistence in weight and height is substantially high, and this persistence is generally driven by time-invariant, unobserved heterogeneity.

Second, for the ECLS-K sample, on the one hand, while persistence in BMI is high over the period spanning kindergarten through eighth grade, predominantly driven by time-invariant, unobserved heterogeneity, focusing separately on weight and height reveals a non-trivial role for state dependence. In other words, in the ECLS-K sample, a seemingly counterintuitive finding is that examining persistence in BMI masks the importance of state dependence in childhood weight and height, at least on average (as evidenced by results in Table 3). For instance, while over 80% and 64% of the persistence in weight and height (on average), respectively, are attributable to state dependence, nearly 80% of the persistence in BMI (on average), is explained by time-invariant heterogeneity. Nonetheless, it is worth reiterating that persistence in the upper tail of the distribution of height and BMI is almost entirely driven by time-invariant, unobserved heterogeneity. Thus, policy-makers seeking to address the childhood obesity epidemic can potentially do more if the early childhood interventions also alter relevant, time-invariant, unobserved attributes. For the ECLS-B sample, on the other hand, persistence in weight and height is predominantly driven by time-invariant, unobserved heterogeneity; state dependence plays very little role.

Third, time-invariant, observed heterogeneity plays little role in both samples, albeit its role is slightly more pronounced in the ECLS-B for reasons mentioned in the previous section. In the ECLS-K sample, children who are male or black, rural or northeast residents, non-native English speakers, had a high birthweight, and have a mother with low education, a low age at first birth, or who participated in the labor force during the child's infancy tend to have higher BMI [as evidenced by inspection of the estimation results of (6)]. In the ECLS-B sample, while the associations between birthweight, gestation age, maternal height and weight, and single versus multiple birth and fetal development are not strong, perhaps due to the complexities involving these relationships that are only currently beginning to be understood in the medical literature, these controls do play a small role in explaining persistence. Nonetheless,

[14] The full regressions for $\widehat{\alpha}$ are available upon request from the authors.

the primary determinants of fetal and infant development that may be critical in placing children on an "obesity trajectory" remain unobserved, even in the ECLS-B.

Lastly, a noteworthy finding that persists across the two samples is the almost negligible effect of time-varying—observed and unobserved—factors on persistence in childhood anthropometric measures. This result is noteworthy in light of the lengthy vector of time-varying, observed factors in both samples as well as the time-varying, unobserved attributes captured by the idiosyncratic error terms in the regression specifications.

The foregoing discussion naturally raises questions or concerns about the attributes captured by or reflected in $\eta$ (i.e., the time-invariant, unobserved heterogeneity), at the minimum. From the analysis presented here, one can only conclude that they are not contained in our set of covariates available in the ECLS-K and ECLS-B samples and they do not at least vary during the primary school years. Nonetheless, prior literature posits some possibilities—prenatal attributes such as maternal BMI, maternal weight gain, maternal smoking, periods of undernutrition during pregnancy, gestational diabetes requiring insulin, and postnatal attributes such as breastfeeding and transitions to solid foods. While we do control for birthweight in both analyses, birthweight alone is not a sufficient proxy for these early influences on fetal development as noted earlier.

Furthermore, a related line of research has documented genetic associations with BMI over the life cycle. For instance, variants of the fat mass and obesity-associated (FTO) gene was associated with increased odds of being overweight or obese in adults and children (Frayling et al. 2007; Dina et al. 2007). Fawcett and Barroso (2010) review studies documenting associations of single nucleotide polymorphisms (SNPs) in the FTO gene region with BMI in multiple populations as well as approaches aimed at understanding other obesity-associated genes. Genetic abnormalities have also been found to be associated with obesity (e.g., Asai et al. 2013; Ramachandrappa et al. 2013). An active area of research is deepening our understanding of obesity-associated genes and the mechanisms or pathways through which they impact BMI with the hope of developing therapeutic interventions.

Recent research finds that genetic associations with BMI may vary by age and cohort. Using longitudinal data from the Framingham Heart Study, Rosenquist et al. (2015) find that the association between the rs993609 variant of the FTO gene and BMI varies by birth cohorts and across time. Walter et al. (2016) also report that a polygenic risk score (GRS–BMI) had a greater association with BMI with more recent birth cohorts. Other studies have found similar genetic associations with BMI that reflect cohort effects (see, e.g., Rokholm et al. (2011); Demerath et al. (2013); Guo et al. (2015)).

Moreover, genetic associations with BMI may be moderated by environment factors. While prior evidence suggests that time-invariant, environmental factors *per se* are not likely to play a significant role, gene–environment interactions may play a non-trivial role. For instance, prior studies have linked gene–environment interactions such as physical activity to reductions in the effect of genetic factors (FTO) on BMI (Li et al. 2010; Kilpeläinen et al. 2011; Ahmad et al. 2013). Using the UK Biobank, Young et al. (2016) find novel interactions between FTO and alcohol consumption, deviations from mean sleep duration, and dietary variations. In particular, the authors find that the effect of FTO increases among those who frequently added salt to food as well as with

deviations from mean sleep duration, while it decreases with the frequency of alcohol consumption. Similarly, Rask-Andersen et al. (2017) document gene–environment interaction effects on BMI for several lifestyle factors and conclude while the genetic effects were diminished for the frequency of physical activity and frequency of alcohol intake, the genetic effects were enhanced by lower socioeconomic status as measured by the Townsend deprivation index (TDI).

Depending on the strength or importance of age-specific and cohort-specific genetic associations with BMI, our findings may not apply to other cohorts or period.[15] Moreover, evidence of gene–environment interactions suggests that some common policy prescriptions (e.g., those that encourage activities that diminish the genetic effects) may still influence obesity trajectories without changing the time-invariant, unobserved factors. However, further research is needed to shed light on the causal mechanisms of gene–environment interaction effects to provide guidance on formulating policies to address such effects.

The findings in this paper imply that, while earlier intervention is preferred to later interventions, *only* interventions that alter the crucial, time-invariant, unobserved risk factors captured by $\eta$ are likely to be effective in combating obesity in the long run. Interventions that leave the attributes captured by $\eta$ unaltered are likely to have, at best, minimal short-run effects and little to no long-run effects. This is entirely consistent with the findings in Davis and Gebremariam (2010), where the authors document that community-based interventions designed to combat childhood obesity that were deemed as successful according to the analysis of data collected via randomized control trials did not produce lasting effects. Eventually, children returned to their "natural state" (Davis and Gebremariam 2010, p. 22). The results are also consistent with Figlio et al. (2014) who document constant effects of birthweight (conditional on gestation length) on cognitive outcomes throughout primary school.

The notion that attributes determined at or shortly after birth, and thus time-invariant over the life of an individual, play a dominant role in the evolution of obesity is *consistent* with the strong evidence in economics and elsewhere on the so-called fetal origins hypothesis (see, e.g., Almond and Currie 2011), with quite profound implications.[16] The fetal origins hypothesis, also referred to as the thrifty phenotype hypothesis or Barker's hypothesis (due to Barker's original publication in 1992), posits long-run effects of conditions *in utero* during critical periods of development through "programmed" changes in the physiology and metabolism of individuals (Barker 1997).

Beginning with Barker's work (Barker 1992), there is a strong belief that *in utero* events may determine whether a fetus ends up on an "obesity trajectory." Deckelbaum and Williams (2001, p. 239S) note that "emerging data suggest associations between the influence of maternal and fetal factors during intrauterine growth and growth during the first year of life, on risk of later development of adult obesity and its comorbidities." More recently, Brisbois et al. (2012) state: "Based on recent evidence,

[15] See Lasky-Su et al. (2008) for evidence on the potential importance of age-varying genetic associations. One important avenue for future research is to account for time-varying, unobserved genetic associations using empirical strategies as in Ding and Lehrer (2014).

[16] While the findings in this paper are consistent with the fetal origins hypothesis, we cannot eliminate other possible explanations for what these salient, unobserved attributes entail.

early-life experiences in utero and postnatal influences may induce permanent changes in physiologic function that programme the long-term regulation of energy balance. This subsequently may adversely impact obesity risk in later life."

Factors that may induce such permanent changes in order to set a fetus upon an "obesity trajectory" are the subject of ongoing research. While initial hypotheses focused on undernutrition and oxygen supply, additional factors such as maternal BMI, maternal weight gain, maternal smoking, gestational diabetes requiring insulin, and postnatal characteristics such as breastfeeding and the timing of introduction to solid foods are also found to be important (Dietz 1997; Deckelbaum and Williams 2001; Brisbois et al. 2012).

Other research closely related to the literature on the fetal origins hypothesis focuses on the identification of early-life physical indicators of predisposition to future obesity. Preliminary results suggest that birthweight, length, and gestation age at birth alone are not strong predictors. Instead, there are complex interactions between these measures, along with other measures such as head circumference, that matter. For example, a fetus born prematurely and, as a result, with low birthweight and length is not likely to be at greater risk of future obesity as long as the fetus' measurements are in proportion and within "normal" ranges given its gestation age. On the other hand, a fetus born with disproportionate physical measurements suggests a greater risk of future obesity (Barker 1997; Godfrey and Barker 2001; Brisbois et al. 2012).

## 6 Conclusion

This paper investigates the origins of persistence in early childhood anthropometric outcomes in a dynamic regression framework, which permits the decomposition of persistence into components reflecting state dependence, observed heterogeneity, and unobserved heterogeneity. While childhood and adult obesity remain important public health concerns in the USA, our understanding that interventions earlier in life are likely to have greater impact has deepened over the past two decades. This understanding follows from well-chronicled evidence that obesity is highly persistent; as such, children may become trapped on an "obesity trajectory" early in life. However, little is known about the origins of these correlations in anthropometric measures over the life cycle. Specifically, whether this correlation reflects state dependence, observed heterogeneity, or unobserved heterogeneity is unknown. Moreover, when this persistence in weight status begins—adolescence, early childhood, postnatally, or prenatally—is also unknown.

Better understanding of the dynamics of weight status is crucial for sound policy-making. If weight is highly persistent and the source of this persistence is state dependence, then small (permanent) changes will have large, long-run effects even if the contemporaneous effects are small. However, if persistence is due to biological or environmental factors that are time-invariant, then the only changes that will have long-run effects are those that alter these underlying factors.

The evidence presented in this paper indicates that there is significant persistence in childhood anthropometric measures—weight, height, and BMI—starting during infancy. Moreover, the persistence in BMI is predominantly due to time-invariant

heterogeneity across individuals determined at birth or shortly thereafter, not state dependence. Moreover, little variation in this time-invariant heterogeneity is explained by attributes observed in the data analyzed here.

If the finding that time-invariant, unobserved attributes play a dominant role in the evolution of obesity is correct, then the most effective interventions to curb obesity may need to start *prior* to childbirth. Similar policy prescriptions to consider preventive measures preconception, during pregnancy, and throughout early childhood have been advocated elsewhere (Brisbois et al. 2012; Deckelbaum and Williams 2001). These findings suggest that strategies to reverse the current childhood obesity epidemic may need to start even earlier than previously thought, namely *in utero*. Examples of such measures may include altering institutional rules concerning federal nutrition programs such as SNAP or WIC, or education provided under these programs, as they relate to pregnant women (e.g., Baum 2012).

## Appendix A. Simulation details

1. **Case I** Own $y_{i0}$, own $x_{it}$, replace $\alpha_i = \widehat{\overline{\alpha}}$, set $\varepsilon_{it} = 0$. This eliminates heterogeneity due to time-invariant and time-varying unobserved factors.
   (a) Compute

$$\widehat{y}_{it} = \widehat{\gamma}\widehat{y}_{it-1} + x_{it}\widehat{\beta} + \widehat{\overline{\alpha}}, \quad t = 1, \ldots, T,$$

   where $\widehat{y}_{i0} = y_{i0}$.
   (b) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{\sum_i I(\widehat{y}_{iT} > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)}.$$

2. **Case II** Own $y_{i0}$, own $x_{it}$, draw $\alpha_i \sim F(\alpha)$ where $F(\cdot)$ is the empirical distribution of $\alpha_i$, set $\varepsilon_{it} = 0$. This allows for time-invariant unobserved heterogeneity, but breaks the correlation between $x$ and $\alpha$.
   (a) Draw $\widetilde{\alpha}_i(r) \sim F(\widehat{\alpha})$, $r = 1, \ldots, R$, where $R$ is the number of simulations.
   (b) Compute

$$\widehat{y}_{it}(r) = \widehat{\gamma}\widehat{y}_{it-1}(r) + x_{it}\widehat{\beta} + \widetilde{\alpha}_i(r), \quad t = 1, \ldots, T,$$

   where $\widehat{y}_{i0}(r) = y_{i0}\ \forall r$.
   (c) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{1}{R} \sum_r \left[ \frac{\sum_i I(\widehat{y}_{iT}(r) > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)} \right].$$

3. **Case III** Own $y_{i0}$, own $x_{it}$, own $w_i$, set $\eta_i = 0$, set $\varepsilon_{it} = 0$. This eliminates heterogeneity due to time-invariant and time-varying unobserved factors.

(a) Compute

$$\widehat{y}_{it}(r) = \widehat{\gamma}\widehat{y}_{it-1} + x_{it}\widehat{\beta} + w_i\widehat{\delta}, \quad t = 1, \ldots, T,$$

where $\widehat{y}_{i0} = y_{i0}$.
(b) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{\sum_i I(\widehat{y}_{iT} > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)}.$$

4. **Case IV** Own $y_{i0}$, own $x_{it}$, draw $w_i \sim F(w)$ where $F(\cdot)$ is the empirical distribution of $w_i$, set $\eta_i = 0$, set $\varepsilon_{it} = 0$. This eliminates heterogeneity due to time-invariant and time-varying unobserved factors and breaks the correlation between $x$ and $w$.
(a) Draw $\widetilde{w}_i(r) \sim F(w), r = 1, \ldots, R$, where $R$ is the number of simulations.
(b) Compute

$$\widehat{y}_{it}(r) = \widehat{\gamma}\widehat{y}_{it-1}(r) + x_{it}\widehat{\beta} + \widetilde{w}_i(r)\widehat{\delta}, \quad t = 1, \ldots, T,$$

where $\widehat{y}_{i0}(r) = y_{i0}\ \forall r$.
(c) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{1}{R}\sum_r \left[\frac{\sum_i I(\widehat{y}_{iT}(r) > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)}\right].$$

5. **Case V** Own $y_{i0}$, own $x_{it}$, own $w_i$, draw $\eta_i \sim F(\eta)$ where $F(\cdot)$ is the empirical distribution of $\eta_i$, set $\varepsilon_{it} = 0$. This breaks the correlation between $x$, $w$ and $\eta$ and eliminates heterogeneity due to time-varying unobserved factors.
(a) Draw $\widetilde{\eta}_i(r) \sim F(\widehat{\eta}), r = 1, \ldots, R$, where $R$ is the number of simulations.
(b) Compute

$$\widehat{y}_{it}(r) = \widehat{\gamma}\widehat{y}_{it-1}(r) + x_{it}\widehat{\beta} + w_i\widehat{\delta} + \widetilde{\eta}_i(r), \quad t = 1, \ldots, T,$$

where $\widehat{y}_{i0}(r) = y_{i0}\ \forall r$.
(c) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{1}{R}\sum_r \left[\frac{\sum_i I(\widehat{y}_{iT}(r) > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)}\right].$$

6. **Case VI** Own $y_{i0}$, replace $x_{it} = \overline{x}_t$, own $\alpha_i$, set $\varepsilon_{it} = 0$. The eliminates heterogeneity due to time-varying factors.
(a) Compute

$$\widehat{y}_{it} = \widehat{\gamma}\widehat{y}_{it-1} + \overline{x}_t\widehat{\beta} + \widehat{\alpha}_i, \quad t = 1, \ldots, T,$$

where $\widehat{y}_{i0} = y_{i0}$.

(b) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{\sum_i I(\widehat{y}_{iT} > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)}.$$

7. **Case VII** Own $y_{i0}$, draw $x_{i\cdot} \sim F(x_1, \ldots, x_T)$ where $F(\cdot)$ is the empirical joint distribution of $x_1, \ldots, x_T$, own $\alpha_i$, set $\varepsilon_{it} = 0$. This breaks the correlation between $x$ and $\alpha$ and eliminates heterogeneity due to time-varying unobserved factors.
   (a) Draw $\widetilde{x}_{i\cdot}(r) \sim F(x_1, \ldots, x_T)$, $r = 1, \ldots, R$, where $R$ is the number of simulations.
   (b) Compute

$$\widehat{y}_{it}(r) = \widehat{\gamma}\widehat{y}_{it-1}(r) + \widetilde{x}_{it}(r)\widehat{\beta} + \alpha_i, \quad t = 1, \ldots, T,$$

   where $\widehat{y}_{i0}(r) = y_{i0}\ \forall r$.
   (c) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{1}{R} \sum\nolimits_r \left[ \frac{\sum_i I(\widehat{y}_{iT}(r) > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)} \right].$$

8. **Case VIII** Own $y_{i0}$, own $x_{it}$, own $\alpha_i$, draw $\varepsilon_{i\cdot} \sim F(\varepsilon_1, \ldots, \varepsilon_T)$ where $F(\cdot)$ is the empirical distribution of $\varepsilon_{i\cdot}$. This breaks the correlation between $\alpha$ and $\varepsilon$.
   (a) Draw $\widetilde{\varepsilon}_{i\cdot}(r) \sim F(\widehat{\varepsilon}_1, \ldots, \widehat{\varepsilon}_T)$, $r = 1, \ldots, R$, where $R$ is the number of simulations.
   (b) Compute

$$\widehat{y}_{it}(r) = \widehat{\gamma}\widehat{y}_{it-1}(r) + x_{it}\widehat{\beta} + \widehat{\alpha}_i + \widetilde{\varepsilon}_{it}(r), \quad t = 1, \ldots, T,$$

   where $\widehat{y}_{i0}(r) = y_{i0}\ \forall r$.
   (c) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{1}{R} \sum\nolimits_r \left[ \frac{\sum_i I(\widehat{y}_{iT}(r) > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)} \right].$$

9. **Case IX** Own $y_{i0}$, own $\alpha_i$, draw $x_{i\cdot}, \varepsilon_{i\cdot} \sim F(x_1, \ldots, x_T, \varepsilon_1, \ldots, \varepsilon_T)$ where $F(\cdot)$ is the empirical joint distribution of $x_1, \ldots, x_T, \varepsilon_1, \ldots, \varepsilon_T$. This breaks the correlation between $\alpha$ and $x, \varepsilon$.
   (a) Draw $\widetilde{x}_{i\cdot}(r), \widetilde{\varepsilon}_{i\cdot}(r) \sim F(x_1, \ldots, x_T, \widehat{\varepsilon}_1, \ldots, \widehat{\varepsilon}_T)$, $r = 1, \ldots, R$, where $R$ is the number of simulations.
   (b) Compute

$$\widehat{y}_{it}(r) = \widehat{\gamma}\widehat{y}_{it-1}(r) + \widetilde{x}_{it}(r)\widehat{\beta} + \widehat{\alpha}_i + \widetilde{\varepsilon}_{it}(r), \quad t = 1, \ldots, T,$$

   where $\widehat{y}_{i0}(r) = y_{i0}\ \forall r$.
   (c) Compute

$$\Pr(y_{iT} > y^* | y_{i0} \geq y_0) = \frac{1}{R} \sum\nolimits_r \left[ \frac{\sum_i I(\widehat{y}_{iT}(r) > y^*) I(y_{i0} > y_0)}{\sum_i I(y_{i0} \geq y_0)} \right].$$

## Appendix B. Tables

See Tables 1, 2, 3, 4, 5, and 6.

**Table 1** Summary statistics for ECLS-K sample

| Variable | Final sample | Initial sample |
|---|---|---|
| *Panel I. Time-varying and time-invariant characteristics for wave 1* | | |
| Male (= 1 if male) | 0.505 | 0.517 |
| | (0.499) | (0.499) |
| White (= 1 if white) | 0.615 | 0.509 |
| | (0.487) | (0.500) |
| Black (= 1 if black) | 0.101 | 0.187 |
| | (0.301) | (0.390) |
| Hispanic (= 1 if Hispanic) | 0.172 | 0.178 |
| | (0.377) | (0.383) |
| Asian (= 1 if Asian) | 0.056 | 0.071 |
| | (0.231) | (0.257) |
| Birthweight (ounces) | 119.466 | 117.626 |
| | (23.154) | (23.944) |
| Premature (= 1 if more than 2 weeks premature) | 0.155 | 0.156 |
| | (0.362) | (0.363) |
| Northeast (= 1 if live in northeast) | 0.190 | 0.181 |
| | (0.392) | (0.385) |
| Midwest (= 1 if live in midwest) | 0.282 | 0.222 |
| | (0.450) | (0.415) |
| South (= 1 if live in south) | 0.314 | 0.347 |
| | (0.464) | (0.476) |
| West (= 1 if live in west) | 0.215 | 0.251 |
| | (0.411) | (0.433) |
| Urban | 0.364 | 0.449 |
| | (0.481) | (0.497) |
| Suburb | 0.380 | 0.394 |
| | (0.485) | (0.489) |
| Rural | 0.255 | 0.157 |
| | (0.436) | (0.363) |
| WIC recipient during pregnancy | 0.321 | 0.409 |
| | (0.467) | (0.492) |
| WIC recipient as infant/child | 0.365 | 0.464 |
| | (0.481) | (0.499) |
| AFDC recipient as infant/child | 0.135 | 0.218 |
| | (0.341) | (0.413) |

**Table 1** continued

| Variable | Final sample | Initial sample |
|---|---|---|
| Food stamps recipient as infant/child | 0.209 | 0.302 |
| | (0.406) | (0.459) |
| Mother's AFB | 24.434 | 23.001 |
| | (5.119) | (4.828) |
| Mother's marital status at birth (= 1 if married) | 0.711 | 0.556 |
| | (0.453) | (0.497) |
| Attended full-day kindergarten (= 1 if attended) | 0.499 | 0.536 |
| | (0.5000) | (0.499) |
| Born in the USA | 0.929 | 0.821 |
| | (0.256) | (0.383) |
| Native English speaker | 0.839 | .791 |
| | (0.368) | (0.406) |
| Attended non-parental prekindergarten | 0.799 | 0.738 |
| | (0.400) | (0.439) |
| Mother worked during child's infancy | 0.695 | 0.643 |
| | (0.461) | (0.479) |
| Mother's education less than high school | 0.112 | 0.147 |
| | (0.315) | (0.354) |
| Mother's education is high school/GED | 0.270 | 0.291 |
| | (0.444) | (0.454) |
| Mother's education is some college | 0.313 | 0.288 |
| | (0.464) | (0.453) |
| Mother's education is 4-year college degree | 0.172 | 0.124 |
| | (0.377) | (0.330) |
| Mother's education is more than 4-year college | 0.093 | 0.061 |
| | (0.291) | (0.239) |
| Index of socioeconomic status | 0.098 | −0.059 |
| | (0.793) | (0.766) |
| Poverty (= 1, if household in poverty) | 0.156 | 0.200 |
| | (0.363) | (0.400) |
| Number of children's books in household | 79.106 | 68.409 |
| | (57.007) | (53.194) |
| Household received AFDC | 0.069 | 0.116 |
| | (0.253) | (0.320) |
| Household received food stamps | 0.120 | 0.185 |
| | (.325) | (0.388) |
| Household size | 4.530 | 4.510 |
| | (1.265) | (1.319) |
| Family type (two parents and siblings) | 0.633 | 0.513 |
| | (0.482) | (0.500) |

**Table 1** continued

| Variable | Final sample | Initial sample |
|---|---|---|
| Family type (two parents and no siblings) | 0.085 | 0.080 |
| | (0.280) | (0.272) |
| Family type (one parent and siblings) | 0.108 | 0.153 |
| | (0.310) | (0.360) |
| Family type (one parents and no siblings) | 0.046 | 0.068 |
| | (0.210) | (0.252) |
| Family type (other) | 0.014 | 0.017 |
| | (0.119) | (0.129) |
| Mother works full time | 0.392 | 0.371 |
| | (0.488) | (0.483) |
| Mother works part time | 0.207 | 0.160 |
| | (0.405) | (0.366) |
| Mother does not work | 0.270 | 0.276 |
| | (0.444) | (0.447) |
| No mother in household | 0.027 | 0.033 |
| | (0.161) | (0.180) |
| Child has health insurance | 0.877 | 0.757 |
| | (0.328) | (0.429) |
| Number of hours watched TV in school week | 1.798 | 1.873 |
| | (1.201) | (1.193) |
| Number of hours watched TV during weekend | 2.422 | 2.500 |
| | (1.463) | (1.491) |
| Household has rules regarding watching TV | 0.843 | 0.739 |
| | (0.364) | (0.439) |
| Number of days household eats breakfast together per weekly | 4.542 | 4.362 |
| | (2.414) | (2.285) |
| Number of days household eats dinner together per weekly | 5.702 | 5.739 |
| | (1.719) | (1.633) |
| Percent of minority students in class at school | 35.922 | 43.931 |
| | (32.301) | (32.787) |
| Neighborhood not safe | 0.029 | 0.035 |
| | (0.167) | (0.185) |
| Neighborhood somewhat safe | 0.224 | 0.247 |
| | (0.417) | (0.431) |
| Neighborhood very safe | 0.700 | 0.565 |
| | (0.458) | (0.496) |
| Household never worried about running out of food | 0.840 | 0.710 |
| | (0.366) | (0.453) |

**Table 1** continued

| Variable | Final sample | Initial sample |
|---|---|---|
| *Panel II. Anthropometric outcomes by wave* | | |
| Weight $z$-scores (wave 1) | 0.264 | 0.261 |
| | (1.094) | (1.059) |
| Weight $z$-scores (wave 2) | 0.393 | |
| | (1.080) | |
| Weight $z$-scores (wave 3) | 0.558 | |
| | (1.101) | |
| Weight $z$-scores (wave 4) | 0.609 | |
| | (1.132) | |
| Weight $z$-scores (wave 5) | 0.673 | |
| | (1.118) | |
| Height $z$-scores (wave 1) | 0.182 | 0.126 |
| | (1.026) | (1.030) |
| Height $z$-scores (wave 2) | 0.076 | |
| | (1.016) | |
| Height $z$-scores (wave 3) | 0.127 | |
| | (1.012) | |
| Height $z$-scores (wave 4) | 0.296 | |
| | (1.020) | |
| Height $z$-scores (wave 5) | 0.166 | |
| | (1.077) | |
| BMI $z$-scores (wave 1) | 0.294 | 0.360 |
| | (1.297) | (1.159) |
| BMI $z$-scores (wave 2) | 0.440 | |
| | (1.103) | |
| BMI $z$-scores (wave 3) | 0.613 | |
| | (1.064) | |
| BMI $z$-scores (wave 4) | 0.669 | |
| | (1.081) | |
| BMI $z$-scores (wave 5) | 0.600 | |
| | (1.107) | |
| Number of observations | 9150 | 11,050 |

Summary statistics (means and standard deviations in parentheses) are based on the ECLS-K sample. While we use data from five waves of the ECLS-K, we report summary statistics for the first wave before and after our sample restrictions described in the text. In Panel I, the first column presents summary statistics for the final sample for wave 1 used in the regression analysis after restrictions, while the second column presents those for the initial sample before any restrictions. Panel II presents the summary statistics for the anthropometric measures for all waves before and after sample restrictions. Note that all sample sizes are rounded to the nearest 50 per NCES restricted data regulations for the ECLS-K

**Table 2** Summary statistics for ECLS-B sample

| Variable | Final sample | Initial sample |
|---|---|---|
| *Panel I. Time-varying and time-invariant characteristics for wave 1* | | |
| Male (= 1 if male) | 0.502 | 0.520 |
| | (0.500) | (0.500) |
| White (= 1 if white) | 0.413 | 0.414 |
| | (0.492) | (0.493) |
| Black (= 1 if black) | 0.154 | 0.163 |
| | (0.361) | (0.369) |
| Hispanic (= 1 if Hispanic) | 0.200 | 0.211 |
| | (0.400) | (0.408) |
| Asian (= 1 if Asian) | 0.109 | 0.116 |
| | (0.312) | (0.320) |
| Northeast (= 1 if live in northeast) | 0.137 | 0.171 |
| | (0.344) | (0.376) |
| Midwest (= 1 if live in midwest) | 0.247 | 0.217 |
| | (0.431) | (0.412) |
| South (= 1 if live in south) | 0.344 | 0.351 |
| | (0.475) | (0.477) |
| West (= 1 if live in west) | 0.272 | 0.260 |
| | (0.445) | (0.440) |
| Urban | 0.714 | 0.740 |
| | (0.452) | (0.440) |
| Suburban | 0.130 | 0.115 |
| | (0.336) | (0.320) |
| Rural | 0.155 | 0.145 |
| | (0.362) | (0.352) |
| Low birthweight | 0.252 | 0.303 |
| | (0.434) | (0.459) |
| Normal birthweight | 0.683 | 0.635 |
| | (0.465) | (0.481) |
| High birthweight | 0.065 | 0.062 |
| | (0.246) | (0.241) |
| Premature (= 1 if more than 2 weeks premature) | 0.147 | 0.177 |
| | (0.354) | (0.382) |
| Birth status (singleton) | 0.807 | 0.854 |
| | (0.394) | (0.3531) |
| Birth status (twin) | 0.176 | 0.133 |
| | (0.381) | (0.340) |
| Birth status (higher order) | 0.007 | 0.008 |
| | (0.081) | (0.088) |

**Table 2** continued

| Variable | Final sample | Initial sample |
|---|---|---|
| Mother's weight before pregnancy (in kilograms) | 66.448 | 65.227 |
| | (16.473) | (15.724) |
| Mother's weight gain during pregnancy | 13.998 | 13.713 |
| | (5.802) | (5.698) |
| WIC recipient during pregnancy | 0.400 | 0.423 |
| | (0.490) | (0.494) |
| WIC recipient as infant/child | 0.397 | 0.413 |
| | (0.489) | (0.492) |
| Mother's AFB | 24.186 | 23.498 |
| | (5.862) | (5.485) |
| Child's age (in months) | 10.477 | 10.570 |
| | (1.8740) | (1.884) |
| Food stamps recipient as infant/child | 0.204 | 0.227 |
| | (0.403) | (0.419) |
| Index of socioeconomic status | 0.012 | −0.118 |
| | (0.865) | (0.852) |
| Poverty (= 1, if household in poverty) | 0.241 | 0.277 |
| | (0.428) | (0.448) |
| Household size | 4.478 | 4.434 |
| | (1.485) | (1.564) |
| Child lives with biological mother | 0.991 | 0.989 |
| | (0.094) | (0.102) |
| Child lives with biological father | 0.794 | 0.759 |
| | (0.404) | (0.428) |
| Child does not live with father | 0.192 | 0.227 |
| | (0.394) | (0.419) |
| Mother's current marital status (= 1 if married) | 0.677 | 0.623 |
| | (0.467) | (0.484) |
| Mother's education less than high school | 0.173 | 0.210 |
| | (0.378) | (0.458) |
| Mother's education is high school/GED | 0.255 | 0.298 |
| | (0.436) | (0.458) |
| Mother's education is some college | 0.280 | 0.253 |
| | (0.449) | (0.435) |
| Mother's education is 4-year college degree | 0.172 | 0.144 |
| | (0.378) | (0.351) |
| Mother's education is more than 4-year college | 0.118 | 0.0921 |
| | (0.323) | (0.289) |

**Table 2** continued

| Variable | Final sample | Initial sample |
|---|---|---|
| Child's primary care with parents | 0.491 | 0.510 |
| | (0.499) | (0.500) |
| Child's primary care with relatives | 0.261 | 0.272 |
| | (0.439) | (0.445) |
| Child's primary care with non-relatives | 0.159 | 0.130 |
| | (0.366) | (0.336) |
| Child's primary care with care center | 0.078 | 0.079 |
| | (0.269) | (0.269) |
| Native English speaker | 0.798 | 0.775 |
| | (0.401) | (0.417) |
| Mother's weight (in kilograms) | 71.464 | 70.525 |
| | (17.388) | (15.831) |
| Family type (two parents and siblings) | 0.545 | 0.492 |
| | (0.498) | (0.499) |
| Family type (two parents and no siblings) | 0.257 | 0.275 |
| | (0.437) | (0.447) |
| Family type (one parent and siblings) | 0.102 | 0.118 |
| | (0.303) | (0.323) |
| Family type (one parents and no siblings) | 0.089 | 0.107 |
| | (0.285) | (0.309) |
| Mother works full time | 0.331 | 0.316 |
| | (0.470) | (0.465) |
| Mother works part time | 0.188 | 0.163 |
| | (0.391) | (0.369) |
| Mother does not work | 0.478 | 0.513 |
| | (0.499) | (0.500) |
| No mother in household | 0.001 | 0.002 |
| | (0.033) | (0.041) |
| Household received food stamps | 0.204 | 0.227 |
| | (0.403) | (0.419) |
| Household received welfare | 0.087 | 0.102 |
| | (0.282) | (0.303) |
| Child has health insurance | 0.969 | 0.957 |
| | (0.172) | (0.202) |
| Number of months on medicaid | 3.372 | 3.669 |
| | (4.671) | (4.834) |
| Child received WIC | 0.520 | 0.558 |
| | (0.500) | (0.497) |
| Household never worried about running out of food | 0.559 | 0.589 |
| | (0.496) | (0.492) |

**Table 2** continued

| Variable | Final sample | Initial sample |
|---|---|---|
| *Panel II. Anthropometric outcomes by wave* | | |
| Weight $z$-scores (wave 1) | −0.175 | −0.263 |
| | (1.474) | (1.601) |
| Weight $z$-scores (wave 2) | −0.023 | |
| | (1.369) | |
| Weight $z$-scores (wave 3) | 0.296 | |
| | (1.120) | |
| Weight $z$-scores (wave 4) | 0.316 | |
| | (1.119) | |
| Height (centimeters) (wave 1) | 72.466 | |
| | (4.130) | |
| Height (centimeters) (wave 2) | 85.182 | |
| | (3.676) | |
| Height (centimeters) (wave 3) | 104.434 | |
| | (5.365) | |
| Height (centimeters) (wave 4) | 111.560 | |
| | (5.367) | |
| Number of observations | 5450 | 6950 |

Summary statistics (means and standard deviations in parentheses) are based on the ECLS-B sample. While we use data from four waves of the ECLS-B, we report summary statistics for the first wave before and after our sample restrictions described in the text. In Panel I, the first column presents summary statistics for the final sample for wave 1 used in the regression analysis after restrictions, while the second column presents those for the initial sample before any restrictions. Panel II presents the summary statistics for the anthropometric measures for all waves before and after sample restrictions. Note that all sample sizes are rounded to the nearest 50 per NCES restricted data regulations for the ECLS-B

**Table 3** Dynamic panel data estimation results (ECLS-K)

| | Model specification | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Panel I. Weight z-scores* | | | |
| Lag weight | 0.931* | 0.932* | 0.775* |
| | (0.003) | (0.003) | (0.067) |
| Time-varying covariates | No | Yes | Yes |
| Time-invariant covariates | No | Yes | No |
| Fixed effects | No | No | Yes |
| Observations | 27,470 | 27,470 | 27,470 |
| Underidentification | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| Endogeneity | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| First-stage F-stat | 88,146.9 | 84,006.9 | 269.0 |

**Table 3** continued

| | Model specification | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Panel II. Height z-scores* | | | |
| Lag height | 0.937* | 0.936* | 0.603* |
| | (0.004) | (0.004) | (0.048) |
| Time-varying covariates | No | Yes | Yes |
| Time-invariant covariates | No | Yes | No |
| Fixed effects | No | No | Yes |
| Observations | 27,470 | 27,470 | 27,470 |
| Underidentification | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| Endogeneity | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| First-stage F-stat | 68,631.4 | 64,548.2 | 673.1 |
| *Panel III. BMI z-scores* | | | |
| Lag BMI | 0.912* | 0.911* | 0.217* |
| | (0.004) | (0.005) | (0.015) |
| Time-varying covariates | No | Yes | Yes |
| Time-invariant covariates | No | Yes | No |
| Fixed effects | No | No | Yes |
| Observations | 27,470 | 27,470 | 27,470 |
| Underidentification | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| Endogeneity | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| First-stage F-stat | 8637.8 | 8518.0 | 820.8 |

Robust standard errors in parentheses. Estimation by GMM. Excluded instrument is the dependent variable twice-lagged. Sample sizes rounded to the nearest 10 per NCES restricted data regulations. Sample includes data from fall kindergarten, spring first, spring third, spring fifth, and spring eighth grades. See text for the list of covariates and further details. In an unreported result (available on request from the authors) tests of the hypothesis that $\gamma = 1$ are rejected at the 1 percent level of significance in all regressions
$\ddagger p < 0.10,\ \dagger p < 0.05, * p < 0.01$.

**Table 4** Dynamic simulation results (ECLS-K sample)

| | Weight $z$-scores | | | Height $z$-scores | | | BMI $z$-scores | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | P1 | P2 | P3 | P1 | P2 | P3 | P1 | P2 | P3 |
| *Benchmark* | 0.840 | 0.762 | 0.118 | 0.606 | 0.467 | 0.030 | 0.746 | 0.715 | 0.142 |
| *Panel I* Own Xs, $\varepsilon = 0$, and | | | | | | | | | |
| $\alpha = E[\alpha]$ | 0.753 | 0.134 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $\alpha \sim f(\alpha)$ | 0.576 | 0.396 | 0.213 | 0.235 | 0.092 | 0.144 | 0.347 | 0.179 | 0.346 |
| *Panel II* Own Xs, $\eta = 0$, $\varepsilon = 0$, and | | | | | | | | | |
| $W = W_i$ | 0.727 | 0.246 | 0.006 | 0.006 | 0.000 | 0.000 | 0.070 | 0.000 | 0.044 |
| $W \sim f(W)$ | 0.703 | 0.196 | 0.006 | 0.003 | 0.000 | 0.000 | 0.055 | 0.000 | 0.054 |

**Table 4** continued

| | Weight $z$-scores | | | Height $z$-scores | | | BMI $z$-scores | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | P1 | P2 | P3 | P1 | P2 | P3 | P1 | P2 | P3 |
| *Panel III* Own Xs, own Ws, $\varepsilon = 0$, and | | | | | | | | | |
| $\eta \sim f(\eta)$ | 0.591 | 0.418 | 0.209 | 0.256 | 0.108 | 0.133 | 0.368 | 0.195 | 0.335 |
| *Panel IV* Own $\alpha$, $\varepsilon = 0$, and | | | | | | | | | |
| $X = E[X]$ | 0.844 | 0.780 | 0.116 | 0.669 | 0.523 | 0.019 | 0.797 | 0.792 | 0.105 |
| $X \sim f(X)$ | 0.849 | 0.784 | 0.120 | 0.663 | 0.523 | 0.018 | 0.795 | 0.792 | 0.107 |
| *Panel V* Own Xs, own $\alpha$, and | | | | | | | | | |
| $\varepsilon \sim f(\varepsilon)$ | 0.846 | 0.786 | 0.121 | 0.635 | 0.506 | 0.021 | 0.766 | 0.734 | 0.125 |
| *Panel VI* Own $\alpha$ and | | | | | | | | | |
| $X, \varepsilon \sim f(X, \varepsilon)$ | 0.843 | 0.779 | 0.124 | 0.635 | 0.509 | 0.022 | 0.765 | 0.735 | 0.125 |
| *Panel VII* Own Xs, Own Ws, Own $\alpha$ and | | | | | | | | | |
| $y_{i0} = E[y_{i0}]$ | 0.337 | 0.468 | - | 0.182 | 0.383 | - | 0.358 | 0.189 | - |
| $y_{i0} \sim f(y_{i0})$ | 0.583 | 0.590 | 0.212 | 0.241 | 0.435 | 0.153 | 0.359 | 0.189 | 0.356 |

P1 denotes $\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \geq$ 85th percentile), P2 denotes $\Pr(y_{iT} \geq$ 95th percentile $|y_{i0} \geq$ 95th percentile), and P3 denotes $\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \geq$ 50th percentile). Benchmark case denotes the observed probability in the sample. Simulations (based on 500 draws) are obtained after estimating the dynamic model $y_{it} = \alpha_i + \gamma y_{it} - 1 + X_{it}\beta + \varepsilon_{it}$, where $\alpha_i = w_i\delta + \eta_i$. f() denotes the empirical distribution of the argument. Sample includes data from fall kindergarten, spring first, spring third, spring fifth, and spring eighth grades. See text for the list of covariates and further details. "–" indicates failed convergence

**Table 5** Dynamic panel data estimation results (ECLS-B)

| | Model specification | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Panel I Weight z-scores* | | | |
| Lag weight | 0.873* | 0.870* | 0.124* |
| | (0.010) | (0.012) | (0.013) |
| Time-varying covariates | No | Yes | Yes |
| Time-invariant covariates | No | Yes | No |
| Fixed effects | No | No | Yes |
| Observations | 10,900 | 10,900 | 10,900 |
| Under identification | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| Endogeneity | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| First-stage F-stat | 112.0 | 1398.4 | 1929.3 |
| *Panel II Height z-scores* | | | |
| Lag height | 0.480* | 0.506* | −0.002 |
| | (0.004) | (0.010) | (0.007) |
| Time-varying covariates | No | Yes | Yes |
| Time-invariant covariates | No | Yes | No |

**Table 5** continued

| | Model specification | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Fixed effects | No | No | Yes |
| Observations | 10,900 | 10,900 | 10,900 |
| Underidentification | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| Endogeneity | $p = 0.000$ | $p = 0.000$ | $p = 0.000$ |
| First-stage F-stat | 770.2 | 6940.2 | 10250.2 |

Robust standard errors in parentheses. Estimation by GMM. Excluded instrument is the dependent variable twice-lagged. Sample sizes rounded to the nearest 50 per NCES restricted data regulations. Sample includes data from waves 1–4 in the ECLS-B. See text for the list of covariates and further details. In an unreported result (available on request from the authors) tests of the hypothesis that $\gamma = 1$ are rejected at the 1 percent level of significance in all regressions
$\ddagger p < 0.10,\ \dagger p < 0.05,\ * p < 0.01$.

**Table 6** Dynamic simulation results (ECLS-B sample)

| | Weight $z$-scores | | | Height (cm) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | P1 | P2 | P3 | P1 | P2 | P3 |
| *Benchmark* | 0.541 | 0.366 | 0.431 | 0.600 | 0.365 | 0.179 |
| *Panel I* Own Xs, $\varepsilon = 0$, and | | | | | | |
| $\alpha = E[\alpha]$ | 0.007 | 0.000 | 0.837 | 0.190 | 0.000 | 0.183 |
| $\alpha \sim f(\alpha)$ | 0.255 | 0.118 | 0.617 | 0.380 | 0.092 | 0.383 |
| *Panel II* Own Xs, $\eta = 0$, $\varepsilon = 0$, and | | | | | | |
| $W = W_i$ | 0.169 | 0.031 | 0.613 | 0.396 | 0.020 | 0.148 |
| $W \sim f(W)$ | 0.101 | 0.007 | 0.719 | 0.292 | 0.007 | 0.296 |
| *Panel III* Own Xs, own Ws, $\varepsilon = 0$, and | | | | | | |
| $\eta \sim f(\eta)$ | 0.332 | 0.166 | 0.556 | 0.444 | 0.136 | 0.297 |
| *Panel IV* Own $\alpha$, $\varepsilon = 0$, and | | | | | | |
| $X = E[X]$ | 0.605 | 0.419 | 0.435 | 0.607 | 0.350 | 0.176 |
| $X \sim f(X)$ | 0.596 | 0.416 | 0.440 | 0.597 | 0.353 | 0.204 |
| *Panel V* Own Xs, own $\alpha$, and | | | | | | |
| $\varepsilon \sim f(\varepsilon)$ | 0.581 | 0.405 | 0.431 | 0.597 | 0.304 | 0.158 |
| *Panel VI* Own $\alpha$ and | | | | | | |
| $X, \varepsilon \sim f(X, \varepsilon)$ | 0.582 | 0.413 | 0.441 | 0.589 | 0.357 | 0.214 |
| *Panel VII* Own Xs, Own Ws, Own $\alpha$ and | | | | | | |
| $y_{i0} = E[y_{i0}]$ | 0.240 | 0.115 | – | 0.396 | 0.123 | – |
| $y_{i0} \sim f(y_{i0})$ | 0.240 | 0.117 | 0.612 | 0.404 | 0.127 | 0.386 |

P1 denotes $\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \geq$ 85th percentile), P2 denotes $\Pr(y_{iT} \geq$ 95th percentile $|y_{i0} \geq$ 95th percentile), and P3 denotes $\Pr(y_{iT} \geq$ 85th percentile $|y_{i0} \leq$ 50th percentile). Benchmark case denotes the observed probability in the sample. Simulations (based on 500 draws) are obtained after estimating the dynamic model $y_{it} = \alpha_i + \gamma y_{it} - 1 + X_{it}\beta + \varepsilon_{it}$, where $\alpha_i = w_i\delta + \eta_i$. f() denotes the empirical distribution of the argument. Sample includes data from waves 1–4 of the ECLS-B. See text for the list of covariates and further details. "–" indicates failed convergence

## References


Ahmad S, Rukh G, Varga TV, Ali A, Kurbasic A, Shungin D, Ericson U, Koivula RW, Chu AY, Rose LM et al (2013) Gene× physical activity interactions in obesity: combined analysis of 111,421 individuals of european ancestry. PLoS Genet 9(7):e1003607

Almond D, Currie J (2011) Killing me softly: the fetal origins hypothesis. J Econ Perspect 25(3):153–172

Anderson TW, Hsiao C (1981) Estimation of dynamic models with error components. J Am Stat Assoc 76(375):598–606

Asai M, Ramachandrappa S, Joachim M, Shen Y, Zhang R, Nuthalapati N, Ramanathan V, Strochlic DE, Ferket P, Linhart K, Ho C, Novoselova TV, Garg S, Ridderstrale M, Marcus C, Hirschhorn JN, Keogh JM, O'Rahilly S, Chan LF, Clark AJ, Farooqi SI, Majzoub JA (2013) Loss of function of the melanocortin 2 receptor accessory protein 2 is associated with mammalian obesity. Science 341(6143):275–278

Barker DJ (1997) Maternal nutrition, fetal nutrition, and disease in later life. Nutrition 13(9):807–813

Barker DJP (1992) Fetal and infant origins of adult disease. British Medical Journal Publishing, London

Baum CL (2012) The effects of food stamps on weight gained by expectant mothers. J Popul Econ 25(4):1307–1340

Brisbois T, Farmer A, McCargar L (2012) Early markers of adult obesity: a review. Obes Rev 13(4):347–367

Cawley J, Meyerhoefer C (2012) The medical care costs of obesity: an instrumental variables approach. J Health Econ 31(1):219–230

Davis K, Christoffel KK (1994) Obesity in preschool and school-age children: treatment early and often may be best. Arch Pediatr Adolesc Med 148(12):1257–1261

Davis MM, Gebremariam A (2010) Economic long-term impacts of interventions aimed at preventing or reducing obesity among children, University of Michigan

Deckelbaum RJ, Williams CL (2001) Childhood obesity: the health issue. Obes Res 9(S11):239S–243S

Demerath EW, Choh AC, Johnson W, Curran JE, Lee M, Bellis C, Dyer TD, Czerwinski SA, Blangero J, Towne B (2013) The positive association of obesity variants with adulthood adiposity strengthens over an 80-year period: a gene-by-birth year interaction. Hum Hered 75(2–4):175–185

Dietz WH (1997) Periods of risk in childhood for the development of adult obesity—what do we need to learn? J Nutr 127(9):1884S–1886S

Dietz WH, Gortmaker SL (2001) Preventing obesity in children and adolescents 1. Annu Rev Public Health 22(1):337–353

Dina C, Meyre D, Gallina S, Durand E, Körner A, Jacobson P, Carlsson LM, Kiess W, Vatin V, Lecoeur C et al (2007) Variation in fto contributes to childhood obesity and severe adult obesity. Nat Genet 39(6):724–726

Ding W, Lehrer SF (2014) Understanding the role of time-varying unobserved ability heterogeneity in education production. Econ Educ Rev 40:55–75

Dor A, Ferguson C, Langwith C, Tan E (2010) A heavy burden: the individual costs of being overweight and obese in the United States. Department of Health Policy, School of Public Health and Health Services, The George Washington University, Washington, DC

Eriksson J, Forsén T, Tuomilehto J, Osmond C, Barker D (2001) Size at birth, childhood growth and obesity in adult life. Int J Obes Relat Metab Disord 25(5):735–740

Fawcett KA, Barroso I (2010) The genetics of obesity: FTO leads the way. Trends Genet 26(6):266–274

Figlio D, Guryan J, Karbownik K, Roth J (2014) The effects of poor neonatal health on children's cognitive development. Am Econ Rev 104(12):3921–3955

Finkelstein EA, Fiebelkorn IC, Wang G (2004) State-level estimates of annual medical expenditures attributable to obesity. Obes Res 12(1):18–24

Finkelstein EA, Zuckerman L (2010) The fattening of America: how the economy makes us fat, if it matters, and what to do about it. Wiley, Hoboken

Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW et al (2007) A common variant in the fto gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316(5826):889–894

Freedman DS, Khan LK, Dietz WH, Srinivasan SR, Berenson GS (2001) Relationship of childhood obesity to coronary heart disease risk factors in adulthood: the Bogalusa heart study. Pediatrics 108(3):712–718

Frisvold DE, Giri A (2014) The potential of early childhood education as a successful obesity intervention. In: Brennan VM, Kumanyika SK, Zambrana RE (eds) Obesity interventions in underserved communities: evidence and directions. Johns Hopkins University Press, Baltimore, MD, pp 25–42

Gillman MW, Rifas-Shiman SL, Kleinman K, Oken E, Rich-Edwards JW, Taveras EM (2008) Developmental origins of childhood overweight: potential public health impact. Obesity 16(7):1651–1656
Godfrey KM, Barker DJ (2001) Fetal programming and adult health. Public Health Nutr 4(2b):611–624
Guo G, Liu H, Wang L, Shen H, Hu W (2015) The genome-wide influence on human bmi depends on physical activity, life course, and historical period. Demography 52(5):1651–1670
Halliday TJ (2008) Heterogeneity, state dependence and health. Econom J 11(3):499–516
Ham JC, Iorio D, Sovinsky M (2013) Caught in the bulimic trap? Persistence and state dependence of bulimia among young women. J Hum Resour 48(3):736–767
Hornstein AS, Greene WH (2012) Usage of an estimated coefficient as a dependent variable. Econ Lett 116(3):316–318
Iughetti L, De Simone M, Verrotti A, Iezzi ML, Predieri B, Bruzzi P, Bernasconi S, Balli F, Bedogni G (2008) Thirty-year persistence of obesity after presentation to a pediatric obesity clinic. Ann Hum Biol 35(4):439–448
Kilpeläinen TO, Qi L, Brage S, Sharp SJ, Sonestedt E, Demerath E, Ahmad T, Mora S, Kaakinen M, Sandholt CH et al (2011) Physical activity attenuates the influence of fto variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. PLoS Med 8(11):e1001116
Kleibergen F, Paap R (2006) Generalized reduced rank tests using the singular value decomposition. J Econom 133(1):97–126
Krebs NF, Jacobson MS, of Pediatrics Committee on Nutrition AA et al (2003) Prevention of pediatric overweight and obesity. Pediatrics 112(2):424
Lamb MM, Dabelea D, Yin X, Ogden LG, Klingensmith GJ, Rewers M, Norris JM (2010) Early-life predictors of higher body mass index in healthy children. Ann Nutr Metab 56(1):16–22
Lasky-Su J, Lyon HN, Emilsson V, Heid IM, Molony C, Raby BA, Lazarus R, Klanderman B, Soto-Quiros ME, Avila L et al (2008) On the replication of genetic associations: timing can be everything!. Am J Hum Genet 82(4):849–858
Li S, Zhao JH, Luan J, Ekelund U, Luben RN, Khaw K-T, Wareham NJ, Loos RJ (2010) Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from epic-norfolk prospective population study. PLoS Med 7(8):e1000332
Millimet DL, Tchernis R (2013a) Anthropometric mobility during childhood, Andrew Young School of Policy Studies Research Paper Series (13–11)
Millimet DL, Tchernis R (2013b) Estimation of treatment effects without an exclusion restriction: with an application to the analysis of the school breakfast program. J Appl Econom 28(6):982–1017
Mocan N, Tekin E (2011) Obesity, self-esteem, and wages. In: Grossman N, Mocan N (eds) Economic aspects of obesity. University of Chicago Press, Chicago
Nader PR, O'Brien M, Houts R, Bradley R, Belsky J, Crosnoe R, Friedman S, Mei Z, Susman EJ et al (2006) Identifying risk for obesity in early childhood. Pediatrics 118(3):e594–e601
Ogden CL, Carroll MD, Flegal KM (2008) High body mass index for age among US children and adolescents, 2003–2006. J Am Med Assoc 299(20):2401–2405
Ogden CL, Carroll MD, Kit BK, Flegal KM (2014) Prevalence of childhood and adult obesity in the United States, 2011–2012. J Am Med Assoc 311(8):806–814
Ogden CL, Carroll MD, Lawman HG, Fryar CD, Kruszon-Moran D, Kit BK, Flegal KM (2016) Trends in obesity prevalence among children and adolescents in the united states, 1988–1994 through 2013–2014. Jama 315(21):2292–2299
Ogden CL, Flegal KM, Carroll MD, Johnson CL (2002) Prevalence and trends in overweight among US children and adolescents, 1999–2000. J Am Med Assoc 288(14):1728–1732
Oster E (2014) Unobservable selection and coefficient stability: theory and evidence. J Bus Econ Stat 1–18
Power C, Lake J, Cole TJ (1997) Review: measurement and long-term health risks of child and adolescent fatness. Int J Obes Relat Metab Disord 21(7):507–526
Ramachandrappa S, Raimondo A, Cali AM, Keogh JM, Henning E, Saeed S, Thompson A, Garg S, Bochukova EG, Brage S et al (2013) Rare variants in single-minded 1 (SIM1) are associated with severe obesity. J Clin Investig 123(7):3042
Rask-Andersen M, Karlsson T, Ek WE, Johansson Å (2017) Gene-environment interaction study for bmi reveals interactions between genetic factors and physical activity, alcohol consumption and socioeconomic status. PLoS Genet 13(9):e1006977
Rokholm B, Silventoinen K, Tynelius P, Gamborg M, Sørensen TI, Rasmussen F (2011) Increasing genetic variance of body mass index during the swedish obesity epidemic. PLoS ONE 6(11):e27135

Rosenquist JN, Lehrer SF, OMalley AJ, Zaslavsky AM, Smoller JW, Christakis NA (2015) Cohort of birth modifies the association between fto genotype and bmi. Proc Natl Acad Sci 112(2):354–359
Serdula MK, Ivery D, Coates RJ, Freedman DS, Williamson DF, Byers T (1993) Do obese children become obese adults? A review of the literature. Prev Med 22(2):167–177
Skinner AC, Perrin EM, Skelton JA (2016) Prevalence of obesity and severe obesity in US children, 1999–2014. Obesity 24(5):1116–1123
Taveras EM, Gillman MW, Kleinman K, Rich-Edwards JW, Rifas-Shiman SL (2010) Racial/ethnic differences in early-life risk factors for childhood obesity. Pediatrics 125(4):686–695
Taveras EM, Rifas-Shiman SL, Oken E, Gunderson EP, Gillman MW (2008) Short sleep duration in infancy and risk of childhood overweight. Arch Pediatr Adolesc Med 162(4):305–311
Ulrick SW (2008) Using semi-parametric methods in an analysis of earnings mobility. Econ Journal 11(3):478–498
Van Cleave J, Gortmaker SL, Perrin JM (2010) Dynamics of obesity and chronic health conditions among children and youth. J Am Med Assoc 303(7):623–630
Walpole SC, Prieto-Merino D, Edwards P, Cleland J, Stevens G, Roberts I (2012) The weight of nations: an estimation of adult human biomass. BMC Public Health 12(1):1
Walter S, Mejía-Guevara I, Estrada K, Liu SY, Glymour MM (2016) Association of a genetic risk score with body mass index across different birth cohorts. Jama 316(1):63–69
Whitaker RC, Wright JA, Pepe MS, Seidel KD, Dietz WH (1997) Predicting obesity in young adulthood from childhood and parental obesity. N Engl J Med 337(13):869–873
Young AI, Wauthier F, Donnelly P (2016) Multiple novel gene-by-environment interactions modify the effect of fto variants on body mass index. Nat Commun 7:12724