Who's Affecting Who?
Obesity Peer Effects in NYC Public Schools

Jonathan L. Presler

Saint Louis University, St. Louis, MO 63103
Sinquefield Center for Applied Economic Research

Abstract

Obesity rates continue to climb in the US. Recent work explores social factors that may contribute to these trends, including social norms. Most work in this area focuses on adults and older youth, missing key early years. This paper explores obesity peer effects in New York City schools for students in grades 1-12. Using student fixed effects, I address issues common in the peer effects literature, such as selection and reflection, to estimate causal effects of social norms on student Body Mass Index (BMI). I explore heterogeneity by race, gender, and grade level. Using unconditional quantile regressions, I show that peer effects decline in BMI. This means that students with obesity are less influenced by social norms in BMI than students with healthy weights. My results imply that small social effects in BMI exist, but policymakers hoping to harness social effects to improve BMI should instead view them as an obstacle to be overcome.

JEL codes: I12, I21, D62
Key words: Peer Effects, Obesity, Spillovers, Children, Adolescents

Working Paper 22-04
08-30-2022

Support for this project comes from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01 grant: 1R01HD070739). I would like to thank the NYC Department of Education's Policy Research and Support Group and Office of Food and Nutrition Services for providing data and Armando Taddei, Stephen O'Brien, and Dennis Barrett for their support. For data support, advice, and suggestions, I also thank Meryle Weinstein, Amy Ellen Schwartz, and Joanna Bailey. I thank seminar participants at the AEFP conference for useful comments on previous drafts. The opinions expressed are those of the author and do not represent the views of the NYC Department of Education or NIH.

1 Introduction

In 2018, one in five US youth presented with obesity, a fourfold increase over four decades.[1] Many factors contributed to this obesity epidemic, including declining food prices, dietary changes, and lower levels of physical activity. Recent work explores social factors as well, such as the effect of peers. Peer effects are of particular interest because they have a multiplier effect which may enhance or mitigate other policy interventions. Peers can also play a direct role in health treatments. We see this in smoking, alcohol, and weight loss programs involving peer support networks.

However, causal identification of peer effects is notoriously difficult due to selection, reflection, and correlated effects. Evidence on the existence of body composition peer effects is mixed, and typically focuses on students in high school, college, or adults. This misses early years, which are particularly significant because obesity is difficult to reverse and childhood obesity leads to additional health complications both during childhood and later in life.

This paper uses student's body mass index (BMI) to explore obesity peer effects among students in grade 1-12 in New York City (NYC) public schools. I use reduced form models and draw on methods from the peer effects literature to deal with issues of selection, reflection, and contextual effects. This paper addresses two key questions. First, do BMI peer effects exist? Second, if these effects exist, which students are most affected by their peers, and which peers have the greatest influence? In particular, I explore trends over time to determine whether BMI peer effects are strongest in elementary school or high school. I examine differences by race and gender, and the extent to which homophily by race and gender play a role in these effects. I employ unconditional quantile regression to shed light on difference in effect by own BMI. I find evidence of small but significant social effects. These effects decline in BMI, meaning that students with healthy weights are more influenced by their peers than students with obesity.

1.1 Background

Understanding the existence and extent of social effects is important for policymakers and researchers because the effects create a social multiplier which can enhance or dampen interventions. There are two key areas of research that explore the link between an individual's BMI and those they interact with. The first is the peer effect literature, which asks whether a person's social network affects their bodyweight. The second literature is experimental and looks at individual choices people make, and whether they model behavior after that of the people to whom they are exposed.

The peer effects literature looks at the cumulative effect of these modeling behaviors and has found strong

[1] See Table 1 in Fryar et al. (2020) for a summary of youth overweight, obesity, and severe obesity prevalence using the NHANES.

correlations in weight and BMI for individuals in the same social network (Christakis and Fowler, 2007, Cohen-Cole and Fletcher, 2008, Fowler and Christakis, 2008). Several papers using the Add Health survey find strongest effects among female students (Renna et al., 2008) and the strongest correlations for those most overweight (Halliday and Kwak, 2009). Additional work used random assignment of students to peer groups in college to find evidence of peer effects on fitness and weight gain (Carrell et al., 2011 and Yakusheva et al., 2011). However, there is disagreement about the existence of social effects on body composition. Even recent estimates vary substantially, from statistically significant and near zero (Asirvatham et al., 2018) to very large (Lim and Meer, 2018 and Strombotne et al., 2019).

There appears to be a consensus in experimental work that people's choices around both type and quantity of food consumed are shaped by the behavior of those around them.[2] This work looks at individual food consumption choices and asks whether people are influenced by the behavior of those around them (modeling behavior). There appear to be two underlying reasons for this modeling. First, people look to others to understand appropriate behavior in a situation. Second, people model the behavior of those with whom they identify (or with whom they wish to identify). These effects impact both the choice of the food consumed (healthy or unhealthy choices) and the quantity consumed.

There are several mechanisms through which a person's social network may influence their weight. Below, I discuss norm setting, social identity and stigma, and direct behavioral interactions.

Those with whom a person regularly interacts help define weight norms. Proximity to individuals of lower (higher) BMI may shift student perspectives on what defines a normal weight, causing a healthy-weight student to feel overweight (underweight). Students with a minority of heavy or light classmates may be attracted to or repulsed by these students, and alter behavior accordingly (shining light or bad apple stories).[3] Burke and Heiland (2007) model weight norming behavior in adults. Agents gain utility from consuming food and incur a cost by deviating from the reference weight, which is derived from the group average weight.[4] Burke et al. (2010) looks at population-level trends and shows that social norms in body composition appear to change over time.[5] If norms can be temporally local, it follows that they may be socially or geographically local as well. In fact, we see this in several studies of social norms around bodyweight. Burke et al. (2010) finds that, controlling for own BMI, individuals from groups with higher mean BMI are less likely than individuals from groups with lower mean BMI to self-classify as overweight. Blanchflower

[2] Cruwys et al. (2015) systematically reviews the food consumption modeling literature and shows consensus that modeling effects are real, a conclusion reached in nearly all reviewed papers. More recent work supports this idea as well (Mecheva et al., 2021).

[3] See Hoxby and Weingarth (2005) for a discussion of common peer effects models.

[4] Burke and Heiland (2007) fit their model to data on individuals' weight, metabolic rate, and desired weight. The reference weight is some fraction of group average weight, as they find being thinner than average is the norm that is valued. They include changes in food prices over time, and link the declines in food costs to rising obesity rates.

[5] They show that controlling for BMI (and a variety of demographics), adults were more likely to consider themselves overweight in the early 1990s than in the early 2000s. Their findings suggest that bodyweight norms change over time.

et al. (2009) looks at differences in education levels and find that highly educated people self-identify as overweight more frequently than lower educated people at the same BMI.

This idea of socially local norms is fundamental to the ways a person's group identity may impact their decisions around bodyweight, food consumption, and fitness. We see this as an important mechanism in both the peer effects and modeling literatures. In the peer effects literature, effects are often most pronounced for individuals who share demographic traits (ex: Arcidiacono and Nicholson, 2005, Horrace et al., 2020). Much experimental work suggests that modeling behavior is stronger when individuals are similar along several dimensions, such as sex, weight, and age (Johnston, 2002, Hermans et al., 2010, McFerran et al., 2010, Cruwys et al., 2015).

Modeling may be enhanced by a higher perception of shared characteristics and identity. Interestingly, students are less likely to model group behavior if they already feel socially accepted (Robinson et al., 2011). This suggests that the primary purpose of modeling behavior may be to prove or reinforce group identity. When an individual does not feel the need to prove their group identity, they have more freedom to buck group trends. Differing from BMI norms may incur social stigma (Burke and Heiland, 2007), although this is unlikely to be a healthy tool for incentivizing weight loss (Burke and Heiland, 2018).

Additionally, there are direct behavioral interactions between individuals that may contribute to BMI spread within a social network. These include more direct knowledge or preference transfers. For example, students may directly share food with their peers - or simply share food preferences, which their peers may incorporate into their own diets. Additionally, as student eating habits are shared, students may be made aware of new foods or they may learn which foods are socially acceptable. Behaviors do not have to relate to food intake, and may relate to physical activity, such as exercise habits, sports participation, or recess activities. People may take social queues about fitness to learn how to do an activity, what appropriate physical activity looks like, and in order to receive positive feedback from engaging in the behavior (McNeill et al., 2006). In Section 4.4 I explore fitness spillovers.

To fix ideas, this paper considers norm setting as the primary mechanism of interest through which school body composition peer effects work. My model focuses on large group averages (school-grade cohorts), such that these effects are likely due to norms within this level. It is unlikely that students directly interact with all students in their cohort, but the BMI norms a student observes in their cohort defines what is normal at a local and relevant level. Students in a school-grade cohort are of similar age and body type – likely making up a large portion of age-level peers a student is exposed to. Most students share demographic similarities with the other members of their school-grade cohort, plausibly making this a more relevant group of age-mates for norm setting than media portrayals. A natural way to explore these norms is to look at within-student effects to see how student behavior changes in response to large-group norms. While I use the idea of social norms to ground the discussion of social spillovers, I cannot exclude the possibility that my estimates pick up some more direct interaction effects (such as food sharing).[6]

2 Data and Sample

This paper uses student-level longitudinal administrative data on student academic performance, socio-demographic characteristics, and school codes, which I link to student health and fitness data.

Student-level demographic data comes from the New York City Department of Education (NYCDOE). These data include socio-demographic characteristics such as gender, ethnicity, grade, residential zip code, an indicator of eligibility for free or reduced-price lunch, whether a student is an English language learner, and student with disability status. Unlike standardized tests, which are only recorded in some grades, NYC records student BMI for all students grades K-12, allowing me to explore within-student peer effects. I compile a panel for academic years 2010-2018.

Student health and fitness data comes from the NYCDOE Fitnessgram program. Data include student weight and height measures, which I use to calculate student BMI. Importantly, I know the age of students (in months) at the time the weight and height measures are taken. Additionally, for students in fourth grade and above, the Fitnessgram program tracks a variety of standardized fitness measures designed to assess student flexibility, strength, and endurance.

I use the CDC's recommended cutoffs for underweight, overweight, obese, and severely obese in order to classify student BMI. My measure of BMI is a z-score, normalized to the CDC's national percentiles from 2000 growth charts (CDC, 2007). This z-score is calculated using the 2000 CDC growth charts, adjusting for the student's age (measured to the nearest month) and gender.[7] Throughout this paper, I refer to student body composition simply as BMI.

My estimation sample consists of students in grades 1-12 NYC public, non-charter schools during the 2011-2018 academic years (2010 measures are used as a lag). I exclude students who do not have both a current year measure of BMI (17.69%) and a lag measure of BMI (9.64 %). I also exclude students in school-grade-cohorts with less than 20 students who meet these criteria (0.37 %). I lose an additional 5.07 % of my sample because these observations are singletons within fixed effects (over 99.9 % of this is due to the set of student fixed effects). Table 1 shows summary statistics for the estimation sample. Coverage

[6] Detection of more direct interaction effects would be better served by a more targeted model, such as interaction-weighted friendship networks (Presler, 2020). Without appropriately weighting or otherwise refining the network, the network systematically overweights the influence of peers with whom a student rarely interacts (such as students in another classroom) and underweights the influence of socially proximate peers (such as close friends).

[7] Using 2000 CDC growth charts standardize comparisons of BMI cutoffs across time. In addition, I follow CDC and WHO guidelines (CDC, 2014) for removing "biologically implausible values" and define any that meet these criteria as missing observations (0.29%).

is always highest in middle school and lowest among high school students. The largest drop in my sample occurs because some students do not have a BMI measure (either current or lag), so in section 5 I check whether coverage predicts student BMI.

2.1 Fitnessgram

The Fitnessgram is an annual health and fitness assessment designed for students grades K-12 in order to determine whether students display a health risk. This is a national program in which school districts across the country participate. In NYC, the Office of School Health collects this data, and the assessment is designed with benchmarking criteria rather than percentiles in order to measure student health against objective criteria. While the Fitnessgram measures five areas of student health, my focus in this paper is on BMI.[8] Data reported are student height and weight, which I use to calculate student BMI. This height and weight assessment is collected for most students grade K-12 in NYC public schools, with the highest coverage of these measures occurring in middle school, and the lowest coverage in high schools.

3 Methodology

3.1 Identification

My objective is to identify the effects of peers' body composition on a students' own body composition. Estimating social externalities is difficult due to a number of well-documented challenges. These include the reflection problem, correlated effects, and selection. Below, I discuss each of these challenges in turn, and how I address them.

First is the reflection problem (Manski, 1993). If a student is affected by a group, and the student also affects that group, then it becomes difficult to disentangle the effect each has on one another due to this feedback loop. In order to mitigate the reflection problem, I use lagged peer outcomes rather than contemporaneous outcomes, as in Sacerdote (2001), Vigdor and Nechyba (2007), Carrell et al. (2009), and McVicar et al. (2018). Because the measures are not contemporaneous, this breaks the feedback loop caused by the behavior of a student affecting their peer's behavior in turn reflecting back to own behavior.

Second, students may experience correlated effects (Manski, 1993), which confound estimates of peer effects. Students who share a classroom or school are subject to many of the same teacher, environment, and treatment shocks due to the shared experience required by being one another's peers. This is not social, and it is possible that observed correlations in outcome may be the result of external treatment and not due to

[8] In Section 4.4 I also explore the effects of peer fitness on BMI. In Appendix A, I discuss in greater detail Fitnessgram's measures of physical activity.

student interaction. Following work using similar methodology (Hoxby, 2000, Vardardottir, 2015, Tonello, 2016, and Carrell et al., 2018), I use a series of fixed effects in an attempt to isolate the variation of interest. For example, school fixed effects remove the correlated effect associated with student environment, cafeteria, and neighborhood.

Third, students often select into their peer groups. As with correlated effects, this is another explanation for observed correlation between the performances of students and their peers, making it difficult for the researcher to understand the causal direction. Do peers have similar outcomes because they are influencing one another, or do they elect to be peers because they have similar outcomes? As pointed out in the discussion of correlated effects, my model explores within-school effects. Within a school, students do not choose which cohort of students to which they will belong. By looking at cohorts (as in Hoxby, 2000, Tonello, 2016, and Carrell et al., 2018), I avoid the issue of students selecting into their peer group.

To summarize my approach, I exploit within-student variation in the body composition of same-grade peers. In order to isolate the variation of interest, I use a series of fixed effects to control for common shocks and correlated effects. The idea is to isolate idiosyncratic variation in grade-mate composition within schools, similar to previous work (Hoxby, 2000, Tonello, 2016, and Carrell et al., 2018).[9] To illustrate, my approach asks whether students have a higher BMI if they attend a school-grade with more overweight students than school-grades with more healthy-weight peers.

3.2 Model

The model I estimate uses within-student variation in BMI that comes from exposure to an average peer BMI within the school-grade. When I construct this average classmate BMI, I use a leave-out-mean (LOM) measure and do not include the student's own BMI, as I am interested in the effect of the student's peers (not their own effect). I include this in a simple regression as follows:

$$y_{isgt} = \lambda \bar{y}_{-i,sg,t-1} + \beta X_{isgt} + \bar{Z}_{-i,sgt} + \eta_i + \theta_s + \sigma_g + \psi_t + \gamma_r + \varepsilon_{isgt} \tag{3.1}$$

where y_{icsg} is the outcome for student i in school s, grade g, and year t is classroom. The class average is $\bar{y}_{-i,sg,t-1}$, which is the lagged outcome for all students in the school-grade cohort with i, but not including student i (it is a leave-out mean). The model also includes student controls in the form of time variant student controls X_{isgt} and a student fixed effect η_i. Time varying controls include student age at the time of the Fitnessgram evaluation and its square. I also include cohort characteristic controls $\bar{Z}_{-i,sgt}$, which

[9] I deviate from much of this work by looking at variation within-student rather than within-institution. This allows my model to bridge school types, following the same student as they move between elementary, middle, and high schools. This allows me to better compare peer effects between grade levels. In Section 5, I show that similar baseline models using within-institution variation rather than individual fixed effects produce similar results.

includes leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status. I also control for correlated effects due to school, grade, year, and residential zip code by including the respective fixed effects θ_s, σ_g, ψ_t, and γ_r. ε_{isgt} is the error term. Because the error may be correlated across students within the same school-grade, standard errors are clustered at this level.

It is important to make sure that the coefficient of interest λ is not confounded by shared treatments or shocks faced by students and their peers. These confounders are most likely to affect estimates of peer effects if the estimates are contemporaneous (Lyle, 2007). This is intuitive, as outcomes measured simultaneously are subject to the peer effects of interest, as well as any common treatment shared by the individuals. I break this connection by using lagged peer outcomes (recall that this also addresses the reflection problem). Thus treatments of concern must be common between a student's current year and their peers' previous year - the treatment must span multiple years. Treatments that meet this requirement include school level treatments and neighborhood treatments, which are captured by fixed effects at the school and residential zip code levels.

4 Results

I begin by building to my preferred specification in Table 2, starting with a simple model that includes only student FE and age controls (Model 1). We see that current year BMI is correlated with peer lagged BMI, but it is unclear how much of this is due to sorting into schools, differences by years, and age. Adding grade and year fixed effects (Model 2) does little to change the estimates, but once a school fixed effect is included in the model (Model 3) we see a large drop in the estimate from 0.183 to 0.085. There is little sensitivity to the inclusion of student residential zip code fixed effects (Model 4) and peer characteristics[10] (Model 5). The change in estimated peer effect from Model 2 to Model 3 illustrates the source of selection. BMI is associated with poverty and other characteristics around which families sort into neighborhoods and schools. The stability of the estimate to both neighborhood fixed effects (270 residential zip codes) and peer characteristics provides evidence that isolating within-school variation in student BMI is key for causal estimates of peer effects which are not the result of selection or correlated effects. Model 5 is my preferred specification, and I build on this model in the rest of the paper.

The estimated effect of peer BMI on own BMI remains stable at 0.085. Recall that these are z-scores, so if all other students in a cohort face a shock or treatment that increases their BMI by one standard deviation, own BMI is expected to increase by 0.085 standard deviations. There are two layers of abstraction which may make understanding the magnitude of these effects difficult. First, these are BMI z-scores. This allows

[10]Cohort characteristics include the percent of school-grade-year peers who are black, white, or Asian/other. The percent of peers who are Hispanic are the reference group, as Hispanic students are the modal student in NYC schools. I also include the percent of peers who are female, eligible for free or reduced price lunch, designated as English language learners, and students with disabilities.

comparison across ages, but the absolute change in BMI is smaller for younger students than older students. Second, BMI units may be less familiar than other units such as pounds. This abstraction is necessary to measure body composition, but the peer mechanism affects BMI through student weight. Figure 1 converts a 0.085 standard deviation effect in own BMI to pounds for each age in my sample. I make two simplifications in presenting this figure. First, I bin ages. In my data, age is measured in months, but I group students by their age in years.[11] Second, I fix height at the average height for each age bin. This means that an average-height six-year-old girl would gain 0.42 pounds from a 1 standard deviation increase in peer weight, and an average-height six-year-old boy would gain 0.38 pounds. This increases with student age, such that an average-height eleven-year-old girl would gain 1.13 pounds from a 1 standard deviation increase in peer weight, and an average-height six-year-old boy would gain 0.97 pounds. At age 15, the effect is larger for boys (1.53 pounds) than for girls(1.51 pounds), and this continues as students grow older. The magnitude of this effect suggests that peers are a meaningful influence on weight, but for the average student, these effects are likely dwarfed by other causes of obesity (such as poverty). However, social spillovers provide a multiplier effect for interventions which encourage healthy weight for students.

4.1 Nonlinear Effects

Understanding the extent to which linear effects exist is important to better understand the spillover effects from targeted health interventions. However, the existence of nonlinear effects means we have the ability to improve welfare through Pareto improvements. The effect of moving a student from one group to another is symmetric if peer effects are linear. In order to test for the existence of nonlinear peer effects in BMI, I add an additional network to the model in Equation 3.1. The new model becomes:

$$y_{isgt} = \lambda_1 \bar{y}_{-i,sg,t-1} + \lambda_2 \bar{\delta}_{-i,sg,t-1} + \beta X_{isgt} + \eta_i + \theta_s + \sigma_g + \psi_t + \gamma_r + \varepsilon_{isgt} \tag{4.1}$$

The additional term $\bar{\delta}_{-i,sg,t-1}$ is the proportion of school-grade-year peers who have a BMI meeting the categorical threshold of overweight, obese, severely obese, or underweight. Thus the model tests the effect, beyond average BMI, of having peers at different body composition extremes. Categories are exclusive, such that students who are obese are not considered overweight.

Table 3 reports results for these nonlinear models. Models 1-4 include the proportion of peers who are overweight, obese, severely obese, and underweight (respectively). Model 5 includes all four categories in the same model. I find no evidence of nonlinear effects in peer body composition. This suggests that total welfare cannot be improved by rearranging peer groups. The baseline estimate itself is robust to the inclusion

[11] I do this in the way that people typically talk about age. That is, $a_y = \lfloor a_m/12 \rfloor$, where a_y is age in years, a_m is age in months.

of these networks, and suggests a persistent, causal effect of student peer body composition on own body composition that is linear.

4.2 Heterogeneous Effects

Tables 3.1 and 4.1 suggest a causal effect of about 0.085 standard deviations. School-grade-year cohorts are large, and it is unlikely that all of these peers are equally important for a given student. In fact, it is likely that a subset of a students' peers account for much of the peer effect (Presler, 2020). In this section, I explore heterogeneous effects along three dimensions: ethnicity, gender, and age (grade).

There are several ways to explore these factors. First, I interact average peer BMI with student ethnicity, gender, and grade level. This differentiates the effect of peer influence by student characteristic. This extends Equation 3.1 with an interaction term:

$$y_{isgt} = \lambda \bar{y}_{-i,sg,t-1} \times \mathbb{1}_{i,dt} + \beta X_{isgt} + \eta_i + \theta_s + \sigma_g + \psi_t + \gamma_r + \varepsilon_{isgt} \tag{4.2}$$

where $\mathbb{1}_{i,dt}$ is an indicator for whether student i has characteristic d during year t.

Another way to explore heterogeneous effects and understand which peers are most relevant is to use homophily.[12] Homophily is the idea that students network with peers who share common characteristics. As students are necessarily grouped by grade in order to define school-grade-year cohorts, I explore homophily by ethnicity and gender. To do this, I limit average classmate BMI to students sharing characteristic d (ethnicity or gender). This changes Equation 3.1 to the following:

$$y_{isgt} = \lambda \bar{y}_{-i,sg,t-1,d(i)} + \beta X_{isgt} + \eta_i + \theta_s + \sigma_g + \psi_t + \gamma_r + \varepsilon_{isgt} \tag{4.3}$$

where $d(i)$ indicates that the peer average BMI $\bar{y}$ only takes into account student i's peers if the peers share student i's demographic characteristic ($d(i)$). As an example, this gives us a gender effect for students sharing a gender, but does not differentiate between male and female students. It is useful to give us an idea whether, on average, students find peers of the same gender to be important for social spillovers.

Alternatively, I could differentiate based on characteristic d. This is similar to the previous model, but I do not combine effects into a categorical effect. Rather I separate, for example, the male and female effects.

[12] A common source of homophily explored in the literature is race and ethnicity (Arcidiacono and Nicholson, 2005, Renna et al., 2008, Lavy et al., 2012, Hsieh and Lin, 2017, Ananat et al., 2018, Billings et al., 2019). Horrace et al., 2020 explores the relative importance of several vectors of homophily, finding shared ethnicity and gender to be the most important for elementary school academic performance.

The model I estimate is as follows:

$$y_{isgt} = \sum_{d}^{D} \lambda_d \bar{y}_{-i,sg,t-1,d} \times \mathbb{1}_{i,d} + \beta X_{isgt} + \eta_i + \theta_s + \sigma_g + \psi_t + \gamma_r + \varepsilon_{isgt} \tag{4.4}$$

where $\bar{y}_{-i,sg,t-1,d}$ is the average BMI of students who share characteristic d (in the same school-grade-year cohort). This is multiplied by $\mathbb{1}_{i,d}$, an indicator function for whether student i has characteristic d, such that each student participates exclusively in one demographic network.

Table 4 reports heterogeneous effects by ethnicity.[13] Model 1 interacts the classroom proportion with a student's own ethnicity (Equation 4.2). Notice that the un-interacted term does not appear in the model, so no ethnic group is left out. Students in the Asian/other category are most affected by their peers, with a significant estimate of 0.140. Black students are next, with an estimate of 0.090, followed by Hispanic students (0.069) and white students (0.061).

Model 2 estimates Equation 4.3, finding that the peer effect of students sharing an ethnicity is 0.049. Model 3 breaks this apart by ethnicity, estimating Equation 4.4. Within-ethnicity networks appear strongest for students in the Asian/other group (0.065), followed by black students (0.051), Hispanic students (0.045), and white students (0.033). This ranking matches the interacted estimates in Model 1. Notice that all of these estimates are smaller than the baseline estimate of 0.085 in Table 3.1. This indicates that ethnicity is an important factor for BMI spillovers, but it is not the dominant mechanism through which BMI spillovers occur. Models 4 and 5 confirm this. Model 4 includes both the shared ethnicity network from Model 2, and the full school-grade-year network. The overall peer effect estimate reduces from 0.085 in the baseline specification to 0.072 when the ethnicity homophily network is included in the model. The estimate of the shared ethnicity network reduces from 0.049 in Model 2 to 0.014 in Model 4. It still remains statistically significant and an important mechanism, but it appears much less important than the overall peer BMI composition. Model 5 replaces the overall homophily network with the separated networks. We see a similar story, with the overall peer effect reduced from 0.085 (baseline) to 0.077 (Model 5). Within-ethnicity networks only remain statistically significant for Asian/other and black students.

Models 6 and 7 combine the interaction and homophily effects. Again, homophily effects are dominated by overall (interacted) effects. We see this in Model 6, where the estimate for the shared ethnicity network (0.013) is comparable to the estimate in Model 4 (0.014) - both of which are less than the estimate in Model 2 (0.049). Model 7 is quite interesting. We see the same rank of interaction effects, with Asian/other (0.129), Black (0.079), Hispanic (0.056), and white (0.043). However, the rank of same ethnicity networks

[13] A very small number of students are missing ethnicity information (0.093% of the sample). This explains the difference in sample size when interacting with ethnicity. When calculating group averages, students with missing ethnicity information simply do not figure into these calculations.

are reversed. Students sharing white ethnicity have the strongest within-network effect (0.019), followed by Hispanic students (0.013), black students (0.11), and Asian/other students (0.010).[14] Notice that white students are most affected by their own group while being least affected by their peers as a whole. About 31% of the social effect for white students comes from other white students. That same percentage is much smaller for Asian/other students (7%), Hispanic students (19%), and black students (12%).

Table 5 reports heterogeneous effects by gender. It should be noted that unlike the results presented previously, not all of the gender results can be considered causal. Recall that the BMI z-score is computed using the CDC's 2000 growth charts (CDC, 2007). Part of this calculation takes into account student gender, making its inclusion endogenous. Model 1 interacts the classroom proportion with a student's own gender (Equation 4.2). Male students appear to be more affected by their peers (0.109) than female students (0.062).

Model 2 estimates Equation 4.3. The average peer effect of students sharing a gender is 0.099. Model 3 estimates Equation 4.4, separating the homophilous effect by gender. The within-gender networks for male students (0.113) is stronger than that for female students (0.083). Unlike with ethnicity, the within-gender network appears quite strong. The estimate for the female network (0.083) is similar to the baseline estimate (0.085), and the other estimates are larger. Models 4 and 5 explore this further by including the overall average network. Note that because there are only two gender homophily networks in this model, the overall network can be interpreted as the heterophily network. Model 4 includes both the shared gender network from Model 2, and the full school-grade-year network. The overall peer effect estimate becomes negative (-0.045). Model 5 expands the overall within-gender network to separate within-male and within-female networks. We again see the within-male network is stronger (0.142) than the within-female network (0.114), and the heterophily network remains similar (-0.043) to Model 4.

Models 6 and 7 combine the interaction and homophily models. We see a similar story to the previous models. In Model 6, the estimate for the gender homophily network is strong (0.128) and the estimates for average cohort BMI interacted with indicators for female (-0.059) and male (-0.025) remain negative. In Model 7, the homophily network is separated into female and male networks. The estimate for the within-female network (0.128) is not statistically different from the within-male network (0.127). The interacted effects are nearly identical to Model 6. One thing to notice in these models is the negative effect of the heterophily network. This implies that, for example, a female student's BMI decreases in her male peers' BMI. Conceptually this is difficult to justify, but it is likely due to the endogenous relationship between gender and zBMI.

In this paper, I am able to do something uncommon in the peer effects literature. NYCDOE conducts the

[14] The overall effect on each ethnicity (Model 1) corresponds to the sum of the cohort effect interacted with ethnicity and the within-ethnicity effect in Model 7. For example, the cohort BMI effect interacted with Hispanic (0.056) and the within-Hispanic BMI effect (0.013) total (0.069) matches the cohort BMI effect interacted with Hispanic in Model 1.

Fitnessgram BMI measures for all K-12 students. This means that by using the lagged body composition of a student's peers, I am able to model the peer effect for students in grades 1-12 and compare these estimates to see if peer effects increase or decrease in student age. The theory is ambiguous. For example, younger students may be less set in their ways and more open to influence from others - in particular their peers. On the other hand, younger students may have less freedom and their primary influences may be their parents or other adults (such as teachers). To explore this peer effect over time, I interact average peer BMI with the student's grade level (Equation 4.2).

Figure 2 shows the results of this interaction - first for the full sample (2a) and then with the sample stratified by gender (2b). Figure 2a shows that the peer effect grows over time, with the exception of grade 1. The grade 1 effect is abnormally large and could represent a strong initial norm setting when students are first exposed to a large number of their peers. Starting in grade 2, we see a moderate peer effect which grows over time.[15] Figure 2b partitions the sample by gender. Between grades 2 and 7, the point estimate for female students is larger than for male students. Most of the grade 8 dip we saw in Figure 2a appears to be attributable to female students, and male students may now have a larger peer effect for most of high school. However, the gender differences are not large, and only significantly vary during grades 5 and 6 (with female students being more affected by their peers). The trends this figure suggests is that female students may be more affected by their peers' BMI in earlier grades, while male students may be more affected in later years, with the switch occurring during middle school.

4.3 Quantile Regressions

We have explored who is affected by their peers in terms of demographic characteristics, but it is also of interest to know who is affected based on BMI. I explore this using unconditional quantile regression, following Firpo et al. (2009). This method has been used in the peer effects literature to explore academic spillovers (Huang and Zhu, 2020). Here, I ask whether students with low BMI are more or less affected by peer body composition than students with high BMI.

Table 8 presents the results of unconditional quantile estimates for every ten percentiles of students' own BMI. Notice that the students most affected by their peers are students with the lowest BMI, and the effect decreases monotonically in student BMI.

Quantiles do not correspond directly with body composition category, but we can think of them similarly. All students who are underweight are below the 10th percentile. All students with severely obesity are in the top decile. Students with obesity make up most observations above the 80th percentile. Most overweight students fall between the 70th and the 80th percentiles. This means we can think of normal weight students

[15] The exception to this trend is a dip in the point estimate at grade 8.

being between the 10th and 70th percentiles in BMI.

It is sub-optimal for student's health outcomes if they are either underweight or overweight. These results suggest that peer effects can have a strong effect on underweight students, but have a small effect on students who are obese or severely obese. This may reflect the difficulty of gaining weight compared to losing it. This is one of major concerns for youth overweight and obesity: once students are overweight it is difficult to lose that weight. There is evidence that overweight children as young as five years old have a twenty percent chance of becoming obese within the year, which will likely persist (Cunningham et al., 2014). These results may be discouraging for policymakers hoping to harness body composition spillovers, as we see stronger effects on normal-weight students than the students we would prefer to be most affected by their peers. These findings contrast with estimates found in Brunello et al. (2020), who use peers' genetic disposition towards obesity rather than actualized BMI the previous year and run similar quantile regressions. This suggests that there are other factors (social and otherwise) that affect student weight beyond genetics, and emphasizes the cumulative nature of weight gain. Even for students who are not predisposed to obesity, it is more difficult to lose weight than gain it.

4.4 Fitness Spillovers

When crafting treatments or policy to harness peer effects, it is important to understand the mechanism. I test whether the fitness of a student's peers affect own BMI. To do this, I use the six fitness measures recorded in the Fitnessgram. While students of every grade participate in the health assessment (BMI), students do not begin taking the fitness tests until the fourth grade. Thus this work involves a subset of the sample used in the previous sections, starting with students in the fifth grade (as the models use peer lag performance).

The Fitnessgram records six measures of fitness. Descriptions of the tests and measures are detailed in Appendix A. In my analysis, I use principle-component analysis with a varimax (orthogonal) rotation to consolidate these six measures into two factors of physical competence, which I call fitness and flexibility.

Previous work has shown that an individual's peers affect their own physical fitness performance at the college level (Carrell et al., 2009, Carrell et al., 2011). Indeed, I find similar results for students in NYC (reported in Appendix B). What is of primary interest in this paper is whether peer fitness affects a student's own BMI. This is one of the potential mechanisms through which a student's peers may affect their BMI, and possibly the most compelling mechanism outside norm setting.

Table 7 presents the results of these models. We see a statistically significant effect of both peer flexibility and peer fitness, but the magnitudes of these effects are dwarfed by the effect of peer BMI. Higher peer flexibility in peers may decrease own BMI and higher peer fitness may increase own BMI, but these effects are insignificant relative to peer BMI. Aggregate fitness levels and norms appear to be an unimportant source of spillovers when it comes to student BMI.

5 Robustness Checks

When using reduced form models of peer effects, there is a concern that the models are not actually detecting peer effects, but some spurious correlations - perhaps due to correlated effects not correctly controlled for by the fixed effects. For example, students in the same school-grade cohort may simply be descriptively similar due to some unaccounted sorting (perhaps into neighborhoods or schools). Students travel through school together with many of the same peers, as most peers progress together from year to year. Thus if I am detecting correlated effects, we should expect the time direction to be reversible. In my baseline specification, I test whether the BMI of student i in year 1 is affected by the BMI of i's peers in year 0. Here, I test whether the BMI of student i in year 1 is affected by the BMI of i's peers in year 2. That is, does the future BMI of my peers affect my current year BMI. The results of this test are found in Model 1 of Table 6, and we can see that this model produces a precise zero estimate. This implies that these effects are meaningful peer effects, rather than correlated effects due to shared environmental factors.

Some student Fitnessgram scores are not recorded, either due to lack of participation in the program or due to breakdowns in the recording process. We may be concerned that students in schools with particularly high or low BMI scores are systematically not participating, which would bias our results. In Model 2 of Table 6, I predict cohort coverage percentage using peer BMI and cannot distinguish the results from zero. This suggests that there is no systematic non-participation based on student BMI.

6 Conclusion

Understanding factors relevant to childhood and youth body weight is important for addressing widely observed increases in overweight and obesity levels. There is active debate whether peer effects are consequential for child and youth bodyweight. In this paper, I address issues of selection and reflection and leverage within-student variation to find evidence of small social effects within school-grade cohorts in NYC public schools. The likely mechanism behind my estimates is social norms.

I find that a one standard deviation increase in lagged peer BMI leads to a 0.085 standard deviation increase in own BMI. A standard deviation change in zBMI is roughly equivalent to the change from an average healthy-weight student to overweight,[16] so we can consider the effect of exposure to a cohort of just-

[16] The average healthy-weight student has a zBMI of 0.014. A zBMI of 1.036 classifies a student as overweight.

overweight peers relative to a cohort of healthy-weight peers.[17] For fifth graders, 0.085 standard deviations translate to about a pound.[18] This means that a standard deviation change in the weight of the median fifth grader's peers would be responsible for about 1% of their total weight, or about 5% of the weight difference between overweight and obese.

I also explore a number of nonlinear models and heterogenous effects based on race and age. I provide evidence that students respond to peer zBMI linearly. Any social attraction or repulsion to underweight or overweight peers is dwarfed by the grade weight norm (peer average). Overall, boys are more affected by peer weight than girls. However, this hides some of the heterogeneity, and girls start out more affected by their peers than boys until about the eighth grade, at which point boys are more affected by their peers. I find evidence of homophily in shared sex and ethnicity. Students appear much more affected by same-sex peers in their cohort.[19] The weak same-ethnicity effects suggest that, at least in the context of school-grade cohorts, social norms in weight cross ethnic lines.

Using quantile regressions, I show that peer effects decrease in weight. This is significant, as the heaviest students would likely see the largest health benefits from weight reduction, but they see the smallest effects through the social norms mechanism. Given the small size of these effects, and the mismatch between who we would like to target with treatments, direct interventions aimed at causes of obesity are likely to be most effective. These include policies that improve student's dietary behaviors and choices, as well as poverty alleviation. Social spillovers and weight norms will likely be a small piece of the mechanism through which these interventions will work.

I find evidence of strong peer effects in fitness, but these do not translate to student BMI. This suggests that policies attempting to harness BMI peer effects should target direct factors, such as diet. Indeed, there is much experimental work showing that peers have an impact on individual food choices and consumption levels. In this area, social pressures are also stronger in an unhealthy direction. Evidence suggests it is easier to influence peers to switch from healthy to unhealthy food choices than the reverse (Cruwys et al., 2015, Binder et al., 2019, Mecheva et al., 2021).

As we develop policies to address the steady increase in overweight and obesity rates, it is important to address early years during childhood and youth. Early weight gain is difficult to reverse and carries significant health risks. In this paper, I show that BMI social norms play a small but significant role in student weight. Social effects are strongest for underweight and healthy weight students, and weakest for students with obesity and severe obesity. As we see increases in overweight and obesity rates over time, changing social norms will

[17]The average overweight student has a zBMI of 1.334, so the hypothetical change if we consider average overweight peers is larger than the one proposed here.

[18]For girls, this is just over a pound, and for boys it is just under a pound.

[19]The strong same-sex effects could be a result of the BMI normalization process, which takes student sex into account.

likely push students towards even heavier weights. As a result, policies targeting student weight in schools should view peer effects as an obstacle to overcome, rather than a potentially helpful multiplier. If these policies are successful, and healthy weights again become the norm, it is likely that social effects can again be a helpful tool for maintaining healthy weight among students.

References

Ananat, Elizabeth, Fu Shihe, and Stephen L. Ross (2018). "Race-specific urban wage premia and the black-white wage gap". In: *Journal of Urban Economics* 108.C, pp. 141–153.

Arcidiacono, Peter and Sean Nicholson (Feb. 2005). "Peer effects in medical school". In: *Journal of Public Economics* 89.2-3, pp. 327–350.

Asirvatham, Jebaraj, Michael R. Thomsen, Rodolfo M. Nayga Jr, and Heather L. Rouse (2018). "Do peers affect childhood obesity outcomes? Peer-effect analysis in public schools". In: *Canadian Journal of Economics/Revue canadienne d'économique* 51.1, pp. 216–235.

Biener, Adam, John Cawley, and Chad Meyerhoefer (Jan. 2018). "The Impact of Obesity on Medical Care Costs and Labor Market Outcomes in the US". In: *Clinical Chemistry* 64.1, pp. 108–117.

Billings, Stephen B., David J. Deming, and Stephen L. Ross (Jan. 2019). "Partners in Crime". In: *American Economic Journal: Applied Economics* 11.1, pp. 126–50.

Binder, Alice, Brigitte Naderer, and Jörg Matthes (2019). "Do children's food choices go with the crowd? Effects of majority and minority peer cues shown within an audiovisual cartoon on children's healthy food choice". In: *Social Science & Medicine* 225, pp. 42–50.

Blanchflower, David G., Bert van Landeghem, and Andrew J. Oswald (May 2009). "Imitative Obesity and Relative Utility". In: *Journal of the European Economic Association* 7.2-3, pp. 528–538.

Brunello, Giorgio, Anna Sanz-de-Galdeano, and Anastasia Terskaya (2020). "Not only in my genes: The effects of peers' genotype on obesity". In: *Journal of Health Economics* 72, p. 102349.

Burke, Mary A. and Frank Heiland (2007). "Social Dynamics Of Obesity". In: *Economic Inquiry* 45.3, pp. 571–591.

Burke, Mary A. and Frank W. Heiland (Jan. 2018). "Evolving Societal Norms of Obesity: What Is the Appropriate Response?" In: *JAMA* 319.3, pp. 221–222.

Burke, Mary A., Frank W. Heiland, and Carl M. Nadler (2010). "From "Overweight" to "About Right": Evidence of a Generational Shift in Body Weight Norms". In: *Obesity* 18.6, pp. 1226–1234.

Carrell, Scott E., Richard L. Fullerton, and James E. West (2009). "Does Your Cohort Matter? Measuring Peer Effects in College Achievement". In: *Journal of Labor Economics* 27.3, pp. 439–464.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka (Nov. 2018). "The Long-Run Effects of Disruptive Peers". In: *American Economic Review* 108.11, pp. 3377–3415.

Carrell, Scott E., Mark Hoekstra, and James E. West (2011). "Is poor fitness contagious?: Evidence from randomly assigned friends". In: *Journal of Public Economics* 95.7, pp. 657–663.

Cawley, John (2004). "The Impact of Obesity on Wages". In: *The Journal of Human Resources* 39.2, pp. 451–474.

— (2008). "Contingent valuation analysis of willingness to pay to reduce childhood obesity". In: *Economics & Human Biology* 6.2, pp. 281–292.

CDC (2007). *2000 CDC Growth Charts*. Tech. rep. National Center for Health Statistics at the Centers for Disease Control and Prevention.

— (2014). *Cut-offs to define outliers in the 2000 CDC Growth Charts*. Tech. rep. National Center for Health Statistics at the Centers for Disease Control and Prevention.

Christakis, Nicholas A. and James H. Fowler (2007). "The Spread of Obesity in a Large Social Network over 32 Years". In: *New England Journal of Medicine* 357.4. PMID: 17652652, pp. 370–379.

Cohen-Cole, Ethan and Jason M. Fletcher (2008). "Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic". In: *Journal of Health Economics* 27.5, pp. 1382–1387.

Cruwys, Tegan, Kirsten E. Bevelander, and Roel C.J. Hermans (2015). "Social modeling of eating: A review of when and why social influence affects food intake and choice". In: *Appetite* 86. Social Influences on Eating, pp. 3–18.

Cunningham, Solveig A., Michael R. Kramer, and K.M. Venkat Narayan (2014). “Incidence of Childhood Obesity in the United States”. In: *New England Journal of Medicine* 370.5. PMID: 24476431, pp. 403–411.

Finkelstein, Eric A., Justin G. Trogdon, Joel W. Cohen, and William Dietz (2009). “Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates”. In: *Health Affairs* 28.Supplement 1, w822–w831.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2009). “Unconditional Quantile Regressions”. In: *Econometrica* 77.3, pp. 953–973.

Fowler, J.H. and N.A. Christakis (2008). “Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais”. In: *Journal of Health Economics* 27.5, pp. 1400–1405.

Fryar, CD, MD Carroll, and J Afful (2020). *Prevalence of overweight, obesity, and severe obesity among children and adolescents aged 2–19 years: United States, 1963–1965 through 2017–2018.* NCHS Health E-Stats. CDC.

Halfon, Neal, Kandyce Larson, and Wendy Slusser (2013). “Associations Between Obesity and Comorbid Mental Health, Developmental, and Physical Health Conditions in a Nationally Representative Sample of US Children Aged 10 to 17”. In: *Academic Pediatrics* 13.1, pp. 6–13.

Halliday, Timothy J. and Sally Kwak (July 2009). “Weight gain in adolescents and their peers”. In: *Economics & Human Biology* 7.2, pp. 181–190.

Hermans, Roel C.J., C. Peter Herman, Junilla K. Larsen, and Rutger C.M.E. Engels (2010). “Social modeling effects on snack intake among young men. The role of hunger”. In: *Appetite* 54.2, pp. 378–383.

Horrace, William C., Hyunseok Jung, Jonathan Presler, and Amy Ellen Schwartz (2020). “What makes a classmate a peer? Examining which peers matter in NYC”. Working Paper.

Hoxby, Caroline (Aug. 2000). *Peer Effects in the Classroom: Learning from Gender and Race Variation.* Working Paper 7867. National Bureau of Economic Research.

Hoxby, Caroline and Gretchen Weingarth (2005). *Taking race out of the equation: School reassignment and the structure of peer effects.* Working Paper.

Hsieh, Chih-Sheng and Xu Lin (2017). “Gender and racial peer effects with endogenous network formation”. In: *Regional Science and Urban Economics* 67, pp. 135–147.

Huang, Bin and Rong Zhu (Oct. 2020). “Peer effects of low-ability students in the classroom: evidence from China’s middle schools”. In: *Journal of Population Economics* 33.4, pp. 1343–1380.

Johnston, Lucy (2002). “Behavioral Mimicry and Stigmatization”. In: *Social Cognition* 20.1, pp. 18–35.

Kindblom, Jenny M., Mattias Lorentzon, Asa Hellqvist, Lars Lönn, John Brandberg, Staffan Nilsson, Ensio Norjavaara, and Claes Ohlsson (2009). “BMI Changes During Childhood and Adolescence as Predictors of Amount of Adult Subcutaneous and Visceral Adipose Tissue in Men”. In: *Diabetes* 58.4, pp. 867–874.

Lakdawalla, Darius N., Jayanta Bhattacharya, and Dana P. Goldman (2004). “Are The Young Becoming More Disabled?” In: *Health Affairs* 23.1. PMID: 15002639, pp. 168–176.

Lavy, Victor, Olmo Silva, and Felix Weinhardt (2012). “The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools”. In: *Journal of Labor Economics* 30.2, pp. 367–414.

Lim, Jaegeum and Jonathan Meer (2018). “How do peers influence BMI? Evidence from randomly assigned classrooms in South Korea”. In: *Social Science & Medicine* 197, pp. 17–23.

Ludwig, David S. and Cara B. Ebbeling (Sept. 2001). “Type 2 Diabetes Mellitus in ChildrenPrimary Care and Public Health Considerations”. In: *JAMA* 286.12, pp. 1427–1430.

Lyle, David S (May 2007). “Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point”. In: *The Review of Economics and Statistics* 89.2, pp. 289–299.

Manski, Charles F. (July 1993). “Identification of Endogenous Social Effects: The Reflection Problem”. In: *The Review of Economic Studies* 60.3, pp. 531–542.

McFerran, Brent, Darren W. Dahl, Gavan J. Fitzsimons, and Andrea C. Morales (2010). “Might an overweight waitress make you eat more? How the body type of others is sufficient to alter our food consumption”. In: *Journal of Consumer Psychology* 20.2, pp. 146–151.

McNeill, Lorna Haughton, Matthew W. Kreuter, and S.V. Subramanian (2006). “Social Environment and Physical activity: A review of concepts and evidence”. In: *Social Science & Medicine* 63.4, pp. 1011–1022.

McVicar, Duncan, Julie Moschion, and Chris Ryan (2018). “Achievement effects from new peers: Who matters to whom?” In: *Economics of Education Review* 66, pp. 154–166.

Mecheva, Margarita de Vries, Matthias Rieger, Robert Sparrow, Erfi Prafiantini, and Rina Agustina (2021). “Snacks, nudges and asymmetric peer influence: Evidence from food choice experiments with children in Indonesia”. In: *Journal of Health Economics* 79, p. 102508.

Moro, Andrea, D. Sebastian Tello-Trillo, and Tommaso Tempesti (2018). “The Impact of Obesity on Wages: the Role of Personal Interactions and Job Selection”. In: *working paper*.

Morrison, Katherine M., Sabina Shin, Mark Tarnopolsky, and Valerie H. Taylor (2015). “Association of depression & health related quality of life with body composition in children and youth with obesity”. In: *Journal of Affective Disorders* 172, pp. 18–23.

Must, A. and R. S. Strauss (1999). “Risks and consequences of childhood and adolescent obesity”. In: *International Journal of Obesity* 23.2, S2–S11.

Mustillo, Sarah, Carol Worthman, Alaattin Erkanli, Gordon Keeler, Adrian Angold, and E. Jane Costello (2003). “Obesity and Psychiatric Disorder: Developmental Trajectories”. In: *Pediatrics* 111.4, pp. 851–859.

Park, M. H., C. Falconer, R. M. Viner, and S. Kinra (2012). “The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review”. In: *Obesity Reviews* 13.11, pp. 985–1000.

Presler, Jonathan (2020). *You are who you eat with: Evidence on academic peer effects from school lunch lines*. Working Paper.

Renna, Francesco, Irina B. Grafova, and Nidhi Thakur (2008). “The effect of friends on adolescent body weight”. In: *Economics & Human Biology* 6.3. Symposium on the Economics of Obesity, pp. 377–387.

Robinson, Eric, Talia Tobias, Lucy Shaw, Elyse Freeman, and Suzanne Higgs (2011). “Social matching of food intake and the need for social acceptance”. In: *Appetite* 56.3, pp. 747–752.

Sacerdote, Bruce (May 2001). “Peer Effects with Random Assignment: Results for Dartmouth Roommates*”. In: *The Quarterly Journal of Economics* 116.2, pp. 681–704.

Schwimmer, Jeffrey B., Tasha M. Burwinkle, and James W. Varni (Apr. 2003). “Health-Related Quality of Life of Severely Obese Children and Adolescents”. In: *JAMA* 289.14, pp. 1813–1819.

Serdula, M.K., D. Ivery, R.J. Coates, D.S. Freedman, D.F. Williamson, and T. Byers (1993). “Do Obese Children Become Obese Adults? A Review of the Literature”. In: *Preventive Medicine* 22.2, pp. 167–177.

Strombotne, Kiersten L., Jason M. Fletcher, and Mark J. Schlesinger (2019). “Peer effects of obesity on child body composition”. In: *Economics & Human Biology* 34. New Findings in Economics and Human Biology:A special Issue in Honor of Founding Editor John Komlos, pp. 49–57.

Sutherland, E. Rand (2014). “Linking obesity and asthma”. In: *Annals of the New York Academy of Sciences* 1311.1, pp. 31–41.

Taylor, Erica D., Kelly R. Theim, Margaret C. Mirch, Samareh Ghorbani, Marian Tanofsky-Kraff, Diane C. Adler-Wailes, Sheila Brady, James C. Reynolds, Karim A. Calis, and Jack A. Yanovski (2006). “Orthopedic Complications of Overweight in Children and Adolescents”. In: *Pediatrics* 117.6, pp. 2167–2174.

Taylor, Valerie H., Mary Forhan, Simone N. Vigod, Roger S. McIntyre, and Katherine M. Morrison (2013). “The impact of obesity on quality of life”. In: *Best Practice & Research Clinical Endocrinology & Metabolism* 27.2. Complications of Obesity, pp. 139–146.

Thorpe, Kenneth E., Curtis S. Florence, David H. Howard, and Peter Joski (2004). “The Impact Of Obesity On Rising Medical Spending”. In: *Health Affairs* 23.Suppl1, W4-480-W4–486.

Tonello, Marco (Aug. 2016). “Peer effects of non-native students on natives’ educational outcomes: mechanisms and evidence”. In: *Empirical Economics* 51.1, pp. 383–414.

Trogdon, Justin G., Eric A. Finkelstein, Charles W. Feagan, and Joel W. Cohen (2012). “State- and Payer-Specific Estimates of Annual Medical Expenditures Attributable to Obesity”. In: *Obesity* 20.1, pp. 214–220.

Vardardottir, Arna (2015). “The impact of classroom peers in a streaming system”. In: *Economics of Education Review* 49, pp. 110–128.

Vigdor, Jacob and Thomas Nechyba (2007). “Peer effects in North Carolina public schools”. In: *In Schools and the equal opportunities problem, ed. Ludger Woessmann and Paul.* MIT Press.

Whitlock, Evelyn P., Selvi B. Williams, Rachel Gold, Paula R. Smith, and Scott A. Shipman (2005). “Screening and Interventions for Childhood Overweight: A Summary of Evidence for the US Preventive Services Task Force”. In: *Pediatrics* 116.1, e125–e144.

Yakusheva, Olga, Kandice Kapinos, and Marianne Weiss (2011). “Peer effects and the Freshman 15: Evidence from a natural experiment”. In: *Economics & Human Biology* 9.2, pp. 119–132.

7 Tables

Table 1: Summary Statistics

Variable	Obs	Mean	Std. dev.
Year Observations Per Student	1,108,083	4.559	2.038
zBMI	5,052,142	0.568	1.145
lag zBMI	5,052,142	0.570	1.158
Underweight*	5,052,142	0.036	0.187
Overweight*	5,052,142	0.179	0.384
Obese*	5,052,142	0.137	0.344
Severely Obese*	5,052,142	0.057	0.232
Female	5,052,142	0.504	0.500
Asian/Other	5,047,430	0.198	0.398
Hispanic	5,047,430	0.401	0.490
Black	5,047,430	0.239	0.426
White	5,047,430	0.163	0.369
Flexibility Factor	3,642,536	0.039	0.969
Fitness Factor	3,642,536	0.045	1.007
Year	5,052,142	2014.5	2.2
Grade	5,052,142	6.213	3.339
Age [months]	5,052,142	141.37	41.25
ELL Status	5,052,142	0.114	0.318
SWD Status	5,052,142	0.127	0.333
FRPL Status	5,052,142	0.721	0.448
Cohort Size	51,801	97.53	97.03
N Schools	1,604		

The sample includes all students grades 1-12 in NYC public, non-charter schools during the 2011-2018 school years who have both a BMI measure and a lag BMI measure, and who are in a school-grade-year cohort with at least 20 students.

Table 2: Baseline Reduced Form Regressions

	(1) zbmi	(2) zbmi	(3) zbmi	(4) zbmi	(5) zbmi
Avg Cohort Lagged zBMI	0.179***	0.183***	0.085***	0.085***	0.085***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Age Controls	X	X	X	X	X
Student FE	X	X	X	X	X
Grade FE		X	X	X	X
Year FE		X	X	X	X
School FE			X	X	X
Res Zip Code				X	X
Cohort Characteristics					X
Observations	5,052,142	5,052,142	5,052,142	5,052,142	5,052,142

Each column represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.
***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

Table 3: Nonlinear Models

	(1) zbmi	(2) zbmi	(3) zbmi	(4) zbmi	(5) zbmi
Avg Cohort Lagged zBMI	0.086***	0.083***	0.086***	0.088***	0.094***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)
% Cohort Overweight Last Year	-0.006				-0.001
	(0.018)				(0.023)
% Cohort Obese Last Year		0.018			0.019
		(0.023)			(0.030)
% Cohort Severely Obese Last Year			-0.003		0.001
			(0.013)		(0.017)
% Cohort Underweight Last Year				0.003	-0.004
				(0.053)	(0.052)
Age Controls	X	X	X	X	X
Student FE	X	X	X	X	X
Grade FE	X	X	X	X	X
Year FE	X	X	X	X	X
School FE	X	X	X	X	X
Res Zip Code	X	X	X	X	X
Cohort Characteristics	X	X	X	X	X
Observations	5,052,142	5,052,142	5,052,142	5,052,142	5,052,142

Each column represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.
***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

Table 4: Heterogeneous Effects: Ethnicity

	(1) zbmi	(2) zbmi	(3) zbmi	(4) zbmi	(5) zbmi	(6) zbmi	(7) zbmi
Avg Cohort Lag zBMI X Asian/Other	0.140***					0.127***	0.129***
	(0.011)					(0.011)	(0.011)
Avg Cohort Lag zBMI X Black	0.090***					0.078***	0.079***
	(0.008)					(0.008)	(0.009)
Avg Cohort Lag zBMI X Hispanic	0.069***					0.056***	0.056***
	(0.009)					(0.009)	(0.009)
Avg Cohort Lag zBMI X White	0.061***					0.048***	0.043***
	(0.012)					(0.012)	(0.013)
Avg Cohort Lag zBMI (Same Ethnicity)		0.049***		0.014***		0.013***	
		(0.004)		(0.002)		(0.002)	
Avg Cohort Lag zBMI (Same Ethnicity: Asian/Other)			0.065***		0.036***		0.010**
			(0.005)		(0.004)		(0.004)
Avg Cohort Lag zBMI (Same Ethnicity: Black)			0.051***		0.012***		0.011***
			(0.005)		(0.004)		(0.004)
Avg Cohort Lag zBMI (Same Ethnicity: Hispanic)			0.045***		-0.003		0.013**
			(0.006)		(0.005)		(0.005)
Avg Cohort Lag zBMI (Same Ethnicity: White)			0.033***		0.001		0.019***
			(0.006)		(0.005)		(0.005)
Avg Cohort Lag zBMI				0.072***	0.077***		
				(0.008)	(0.008)		
Age Controls	X	X	X	X	X	X	X
Student FE	X	X	X	X	X	X	X
Grade FE	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X
School FE	X	X	X	X	X	X	X
Res Zip Code	X	X	X	X	X	X	X
Cohort Characteristics	X	X	X	X	X	X	X
Observations	5,047,423	5,052,142	5,052,142	5,052,142	5,052,142	5,047,423	5,047,423

Each column represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.
***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

Table 5: Heterogeneous Effects: Gender

	(1) zbmi	(2) zbmi	(3) zbmi	(4) zbmi	(5) zbmi	(6) zbmi	(7) zbmi
Avg Cohort Lag zBMI X Female	0.062***					-0.059***	-0.059***
	(0.008)					(0.009)	(0.010)
Avg Cohort Lag zBMI X Male	0.109***					-0.025***	-0.024**
	(0.008)					(0.009)	(0.012)
Avg Cohort Lag zBMI (Same Gender)		0.099***		0.131***		0.128***	
		(0.006)		(0.004)		(0.004)	
Avg Cohort Lag zBMI (Same Gender: Female)			0.083***		0.114***		0.128***
			(0.006)		(0.005)		(0.007)
Avg Cohort Lag zBMI (Same Gender: Male)			0.113***		0.142***		0.127***
			(0.006)		(0.005)		(0.008)
Avg Cohort Lag zBMI				-0.045***	-0.043***		
				(0.008)	(0.008)		
Age Controls	X	X	X	X	X	X	X
Student FE	X	X	X	X	X	X	X
Grade FE	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X
School FE	X	X	X	X	X	X	X
Res Zip Code	X	X	X	X	X	X	X
Cohort Characteristics	X	X	X	X	X	X	X
Observations	5,052,142	5,052,142	5,052,142	5,052,142	5,052,142	5,052,142	5,052,142

Each column represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.
***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

Table 6: Placebo Models

	(1) zbmi	(2) fgcoverage
Avg Cohort Lead zBMI	-0.000***	
	(0.000)	
Avg Cohort Lag zBMI		0.001
		(0.002)
Age Controls	X	X
Student FE	X	X
Grade FE	X	X
Year FE	X	X
School FE	X	X
Res Zip Code	X	X
Cohort Characteristics	X	X
Observations	4,996,645	5,052,142

Each column represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.
***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

Table 7: Effects of Cohort Fitness on BMI

	(1) zbmi	(2) zbmi	(3) zbmi
Avg Cohort Lag zBMI	0.074***	0.073***	0.074***
	(0.008)	(0.008)	(0.008)
Avg Cohort Lag Flexibility	-0.006**		-0.006
	(0.003)		(0.006)
Avg Cohort Lag Strength/Endurance	0.007**		0.013**
	(0.003)		(0.006)
% Cohort Fit [Flexibility]		-0.005*	0.001
		(0.003)	(0.008)
% Cohort Fit [Strength/Endurance]		0.004	-0.008
		(0.003)	(0.006)
Age Controls	X	X	X
Student FE	X	X	X
Grade FE	X	X	X
Year FE	X	X	X
School FE	X	X	X
Res Zip Code	X	X	X
Cohort Characteristics	X	X	X
Observations	3,088,370	3,087,805	3,087,805

Each column represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.
***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

Table 8

Quantile	Estimate	Standard Error
q10	0.148***	(0.014)
q20	0.120***	(0.011)
q30	0.106***	(0.009)
q40	0.092***	(0.008)
q50	0.088***	(0.007)
q60	0.079***	(0.007)
q70	0.060***	(0.006)
q80	0.044***	(0.006)
q90	0.026***	(0.005)

Each row represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. As in previous models, all regressions include age controls, student FE, grade FE, year FE, school FE, residential zip code FE, and cohort characteristics. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.

***Significant at the 1% level

**Significant at the 5% level

*Significant at the 10% level

8 Figures

Figure 1: Interpreting 0.085 Standard Deviations as Pounds, by Mean Height for Age

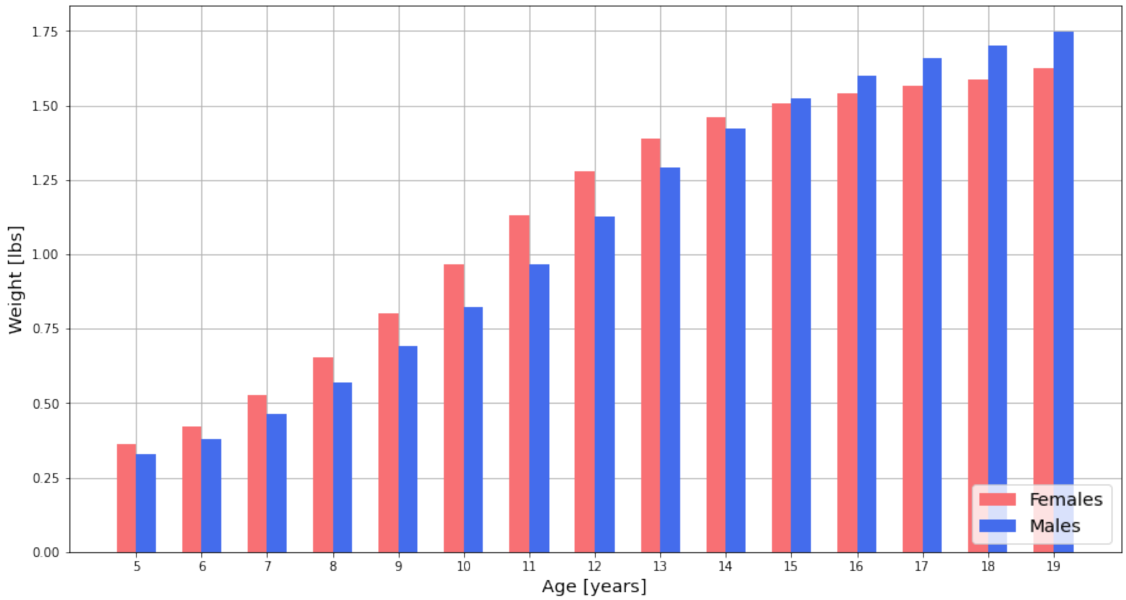
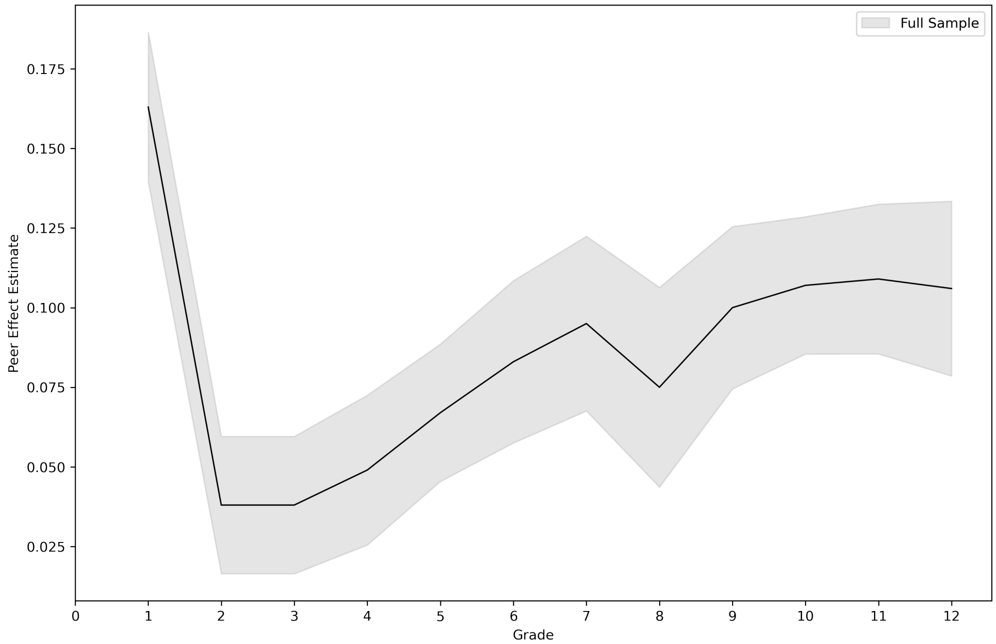
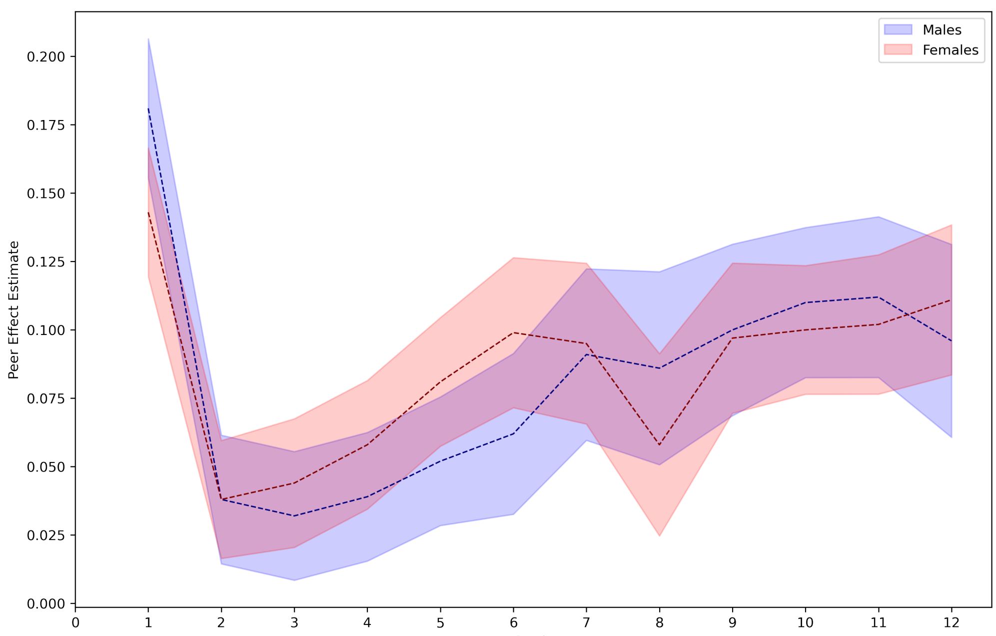


Figure 2: zBMI Peer Effects by Grade



(a) zBMI Peer Effects by Grade, Combined



(b) zBMI Peer Effects by Grade, Stratified by Gender

A Fitnessgram Tests

In NYC, students in all grades participate in the health evaluation that consists of height and weight measurements. These measures are taken by physical education teachers or school nurses - all of whom receive training in order to standardize the process and ensure the accuracy of these measures. Beginning in the fourth grade, students also participate in five physical fitness evaluations. These are the PACER, push-up, curl-up, trunk-lift, and sit-and-reach tests.

The PACER test measures student aerobic capacity by measuring the number of 20 meter laps students can complete at a given pace. The pace is standardized using pre-recorded audio, which progressively gets faster as the test continues. Students must run the twenty meter track between beeps. Once a student misses two beeps, their run ends and their physical education teacher records the number of completed laps.

Similar to the PACER test, the push-up and curl-up tests are conducted using pre-recorded audio which defines the pace at which students complete the exercise. Once students make two pre-defined form errors (which includes things like not maintaining the proper pace or not reaching a 90 degree bend in their elbows for push-ups), their completed exercise reps are recorded.

The sit-and-reach and trunk-lift exercises measure flexibility. The sit-and-reach measures flexibility in the hamstrings, and measures are taken for each leg separately using a box and measuring the distance students are able to reach. The trunk-lift is designed to measure low back health. In it, students lay on their stomachs and raise their back in order to measure how high they are able to bring their chin off the ground while maintaining spinal alignment.

B Fitness Spillovers

Table 9 shows the results of social spillovers for fitness outcomes. As in my BMI specifications, I regress current year fitness results on the cohort's average lag performance. Both current-year and lag measures are normalized z-scores. My estimates range from 0.107 in the sit and reach to 0.192 in the curlup. Notice that the lowest spillovers occur in the flexibility measures, and the largest effects occur in the strength and endurance measures. This is intuitive. Peer encouragement and competition likely extract additional effort from students, and it is likely that effort translates into endurance and strength activities more than flexibility activities. It is also worth noting that although we look at different age groups, my results are strikingly similar to those found in Carrell et al. (2011). They find that one standard deviation increase in high school fitness scores (lag fitness score) increases college fitness by 16.5% of a standard deviation.

C Public and Private Costs of Obesity

This appendix provides a very brief overview of both public and private costs associated with obesity. These include a number of public externalities that make a reduction in obesity rates desirable.

Childhood obesity is a strong predictor of adult obesity (Serdula et al., 1993, Must and Strauss, 1999, Kindblom et al., 2009, and Whitlock et al., 2005). Obesity in adults is linked with a number of health concerns including asthma (Sutherland, 2014), depression and other psychiatric disorders (Mustillo et al., 2003), type 2 diabetes, hyperlipidemia, hypertension, and heart disease (Thorpe et al., 2004 and Park et al., 2012). In children, being overweight has been linked to a number of immediate problems during childhood, including heightened risk of orthopedic, neurological, pulmonary, gasteroenterological, and endocrine conditions, joint problems, allergies (Must and Strauss, 1999, Taylor et al., 2006, and Halfon et al., 2013), type 2 diabetes (Ludwig and Ebbeling, 2001), depression (Morrison et al., 2015 and Halfon et al., 2013), and an overall decrease in quality of life (Taylor et al., 2013 and Schwimmer et al., 2003). Obesity in children leads to problems in adulthood including increased risk of type 2 diabetes, hypertension, heart disease, and mortality (Must and Strauss, 1999 and Park et al., 2012). Obese school children also are more likely to develop ADHD, learning disabilities, developmental delays, grade repetition, and missed school days (Halfon et al., 2013).

In addition to these health concerns, overweight individuals (particularly white females) may see lower wages than healthy-weight counterparts (Cawley, 2004 and Moro et al., 2018), as well as lower chance of being employed at all (Biener et al., 2018).

Problems with overweight and obesity are not limited to private health or wage concerns. Health care costs in the US continue to rise, and a sizable piece of these expenditure increases can be attributed to continuing increases in overweight and obesity (Thorpe et al., 2004, Trogdon et al., 2012, and Biener et al., 2018). These medical costs are borne both privately and publicly. Trogdon et al. (2012) find that between about a quarter and half of these costs are financed publicly through medicare and medicaid (variation by state). In addition, Cawley (2008) estimates that voters are interested in reducing childhood obesity and that the state of New York has a willingness to pay of approximately $690 million to reduce childhood obesity by half in the state (which is higher than current spending levels). Medical expenditures in the US are estimated to be between 6 and 11 percent lower in the absence of obesity (Trogdon et al., 2012 and Biener et al., 2018), with an estimated total cost of obesity in the US of nearly $150 billion in 2008 (Finkelstein et al., 2009). In addition, Lakdawalla et al. (2004) links obesity to higher levels of disability insurance, medicare, and medical costs.

D Appendix Tables and Figures

Table 9: Fitness Spillovers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Avg Cohort Lag zPacer	0.131***									
	(0.012)									
Avg Cohort Lag zPushup		0.156***								
		(0.010)								
Avg Cohort Lag zCurlup			0.192***							
			(0.010)							
Avg Cohort Lag zTrunkLift				0.113***						
				(0.009)						
Avg Cohort Lag zSitreach (L)					0.108***					
					(0.012)					
Avg Cohort Lag zSitreach (R)						0.107***				
						(0.012)				
Avg Cohort Lag Flexibility							0.116***			
							(0.011)			
Avg Cohort Lag Fitness								0.182***		
								(0.009)		
% Cohort Fit [Flexibility]									0.121***	
									(0.009)	
% Cohort Fit [Strength/Endurance]										0.180***
										(0.008)
Age Controls	X	X	X	X	X	X	X	X	X	X
Student FE	X	X	X	X	X	X	X	X	X	X
Grade FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
School FE	X	X	X	X	X	X	X	X	X	X
Res Zip Code	X	X	X	X	X	X	X	X	X	X
Cohort Characteristics	X	X	X	X	X	X	X	X	X	X
Observations	3,074,602	3,078,682	3,078,523	3,082,469	3,083,270	3,083,272	3,044,240	3,044,240	3,043,266	3,043,266

Each column represents a separate regression. Robust standard errors are in parentheses, clustered at the school-grade level. Age controls include student age to the nearest month and its square. Cohort characteristics include leave-out-means for gender, ethnicity, FRPL status, ELL status, and SWD status.
***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level