How Do Peers Influence BMI?
Evidence from Randomly Assigned Classrooms in South Korea

Jaegeum Lim[1] and Jonathan Meer[2,3*]

[1]Korean National Assembly, Seoul, South Korea
[2]Department of Economics, Texas A&M University,
TAMU 4228, College Station, TX 77843, USA
[3]National Bureau of Economic Research, Cambridge, MA, USA

[*]E-mail: jmeer@tamu.edu.

November 8, 2017

Abstract

Obesity among children is an important public health concern, and social networks may play a role in students' habits that increase the likelihood of being overweight. We examine data from South Korean middle schools, where students are randomly assigned to classrooms, and exploit the variation in peer body mass index. We use the number of peers' siblings as an instrument to account for endogeneity concerns and measurement error. Heavier peers increase the likelihood that a student is heavier; there is no spurious correlation for height, which is unlikely to have peer contagion. Public policy that targets obesity can have spillovers through social networks.

Introduction

The proportion of overweight children in OECD countries has risen steadily over the past several decades (OECD, 2015), and estimates of the medical costs of youth obesity are substantial (Biener et al., 2017). Social factors rather than individual characteristics may play a large role, especially for adolescents. However, empirical research on the impacts of peer effects and social contagion on obesity is difficult because of well-known issues such as self-selection, common environmental factors (frequently called contextual effects), and the "reflection problem," in which the direction of influence is unclear (Manski, 1993; Epple and Romano, 2011). Christakis and Fowler (2007) use measured height and weight data in Framingham Heart Study, spanning 32 years, to show that an individual's own chance of becoming obese increases by 57 percent if his or her friend is obese. By examining the correlation of obesity within clusters of social ties, they argue their effect is not due to common factors. However, there is some controversy about the validity of their identification strategy, with others claiming that these correlations are spurious (Cohen-Cole and Fletcher, 2008b). Replicating Christakis and Fowler (2007) with a large sample of adolescents, Cohen-Cole and Fletcher (2008b) show that peer effects on obesity are reduced substantially once controls for contextual effects are included. In further work, they show that similar peer effects can be found in acne, headaches, and height using the same empirical approach, but that these effects also disappear when using a more comprehensive estimation approach (Cohen-Cole and Fletcher, 2008a); Fowler and Christakis (2008) dispute a number of these claims.

Since this work, there have been two general strategies to address these empirical challenges. The instrumental variables approach uses pre-determined peer characteristics or the characteristics of peers' relatives or other friends to provide identification (Trogdon et al., 2008; Mora and Gil, 2013; Cawley et al., 2017). However, these instruments may still be correlated with unobserved peer attributes that cause social network formation. Another approach is to use

institutional settings in which peers are randomly assigned. For example, more-fit peers improve fitness outcomes among cadets who are randomly assigned to squadrons at the United States Air Force Academy (Carrell et al., 2011). Similarly, there is peer influence in weight gain among female college students randomly assigned to dormitory rooms (Yakusheva et al., 2014).

In this paper, we avoid the nonrandom sorting problem by using a unique practice of Korean middle schools: the random assignment of students into a physical homeroom classroom in which they stay with the same classmates throughout the day for an entire school year. Because of the amount of time spent with these peers, they form a useful social network to examine. Our data, drawn from a sample of seventh graders, come with two shortcomings, however. First, they are cross-sectional, meaning that we cannot use lagged observations to better ascertain the direction of influence – that is, the reflection problem. Second, student height and weight are self-reported. While self-reports have generally been found to be reliable, it is possible that students with particularly tall or heavy peers systematically misreport their own height and weight, leading to spurious correlation (Strauss, 1999; Kuczmarski et al., 2001; Lee et al., 2011).

We address these issues by using peers' number of siblings as an instrumental variable for peers' average body mass index (BMI). Numerous studies have found that the number of siblings is correlated with a child's BMI and the likelihood of being obese (Hesketh et al., 2007; Chen and Escarce, 2010; Haugaard et al., 2013; de Oliveira Meller et al., 2015). Mechanisms posited for this correlation include that children with more siblings tend to have higher level of physical activity and lower food intake. Importantly, the number of siblings of the randomly-assigned peers in Korean classrooms cannot directly affect a student's own health. In the Methods Section, we describe this instrument in greater detail and provide evidence for its validity.

Our estimates show that a one unit increase in average peers' BMI increases a student's BMI by 0.83 units. Follow-on data, which unfortunately does not include a sufficient number of peers for a lagged approach as in Christakis and

Fowler, does indicate that seventh grade peers still have a substantial impact on the student's BMI in eighth grade. Our results are robust to the inclusion of a number of student, peer, and teacher characteristics. We further show that our results are driven by peer effects in weight, rather than height; effects on the latter would suggest that our results are tainted by non-classical measurement error or unresolved non-random sorting.

While South Korea has rates of overweight that are among the lowest in the OECD, the proportion of men who are overweight has increased rapidly in recent years (OECD, 2010). Among male children, the proportion of boys who are overweight is actually greater than the OECD average (OECD, 2015). Our results are therefore relevant for other developed countries, and the strength of our identification approach provides important evidence on the impact of social networks on health outcomes.

Methods

Data Set

We use the Gyeonggi Education Panel Study (GEPS2012), which surveyed 4,051 seventh grade students in middle schools in Gyeonggi province (surrounding Seoul, South Korea). Students were sampled with a two-stage cluster sampling design; first, 63 schools were chosen from the population of 624 middle schools in Gyeonggi province. Then, two classrooms were drawn within each school, and all students in the classrooms were surveyed. GEPS also surveyed parents, homeroom teachers, principals, and schools. Each student is linked to their homeroom teacher, who manages the classroom to which the students belong; subject teachers rotate through to present lectures to the same set of students (Lim and Meer, forthcoming).

Students were asked to report their height and weight, with which we construct BMI. BMI has been criticized as sometimes misclassifying an individual as obese or overweight when he or she is muscular, because it cannot distinguish

adipose tissue from muscle, bone, and other lean body mass (Burkhauser and Cawley, 2008). Nevertheless, BMI is a convenient measure, particularly because it rescales to a fairly consistent metric across ages (Bellizzi and Dietz, 1999).

GEPS surveys students in 8th and 9th grade as well, and students once again provide their height and weight. However, since not every classroom in a school is surveyed, data on peers is incomplete in 8th and 9th grades. While we examine the impact of *7th* grade peers on later BMI outcomes, we do not examine contemporaneous peer effects in those later grades.

Starting with 4,051 students, we drop 128 observations that have missing height or weight information. We have a final sample of 3,909 students after dropping 14 additional students who do not have classroom information or whose parents report that the sum of male and female children at home is zero. There is no correlation between attrition from the sample and other student or teacher characteristics. Table 1 reports summary statistics for our sample.

Tests of Random Assignment

Peer Assignment   The random assignment of peers is crucial to our identification strategy. Middle school classrooms in Korea use some form of random assignment to classrooms due to both social norms and government policies (Lim and Meer, forthcoming; Kang, 2007). Students are re-randomized in each year through 9th grade. This is an important feature for examining the impact of peers on later outcomes; if students were somehow sorted into classrooms based on their BMI, the measured impacts of 7th grade peers on 8th and 9th grade BMI might actually reflect this sorting rather than a causal effect.

We also provide empirical evidence to support this point. First, we conduct a series of Pearson's χ^2 tests for the independence of students' assigned classroom and students' characteristics, including gender, number of siblings, father's and mother's education, parents' marital status, as well as whether parents own their own home. Eleven of 378 tests are not available in seventh grade because five schools are single-sex (5 tests) and one school has only one classroom sampled

$(1 \times 6$ tests). Five tests are unavailable for eighth and ninth grades.

Table 2 shows the number of rejections for the null hypothesis of independence in a series of Pearson's χ^2 tests at 1, 5, and 10 percent significance level. For each significance level, the rejection rates are below or near the significance level. Hence, we conclude there is little evidence of nonrandom assignment of students into classroom with respect to student's observable characteristics.

We also test whether the characteristics of classroom peers are correlated with the student's own characteristics. In general, regressing own characteristics on peers' characteristics is not appropriate because of the negative bias that is inherent with random assignment (Guryan et al., 2009). Intuitively, when peers are randomly assigned, the classroom averages for student characteristics will be balanced across the classrooms. If, for example, a student's family income is higher than the classroom mean, then the peers' family income, which excludes that data point, will tend to be lower. This creates a mechanical negative correlation between own and peers' characteristics in spite of random assignment. Monte Carlo studies show that this simple test of correlations over-rejects independence at a very high rate; this issue is resolved by the use of the leave-out mean for the population from which peers are drawn (for example, the school-level rather than classroom-level average family income) (Guryan et al., 2009).

We estimate the following equation to test the independence of characteristics:

$$y_{ics} = \alpha + \beta \frac{1}{n_c - 1} \sum_{j \neq i, j \in c(i)} y_{jcs} + \gamma \frac{1}{n_s - 1} \sum_{j \neq i, j \in s(i)} y_{js} + \lambda_s + \varepsilon_{ics}, \qquad (1)$$

where y_{ics} is student i's characteristic in classroom c in school s. n_c and n_s are the numbers of students in classroom c and school s, respectively, so that the term including y_{jcs} is the mean of the characteristic within the classroom (excluding the student him- or herself). The term including y_{js} is the mean of the characteristic within the school, also excluding the individual student. λ_s

7

are school fixed effects to account for the fact that random assignment occurs within a school and ε_{ics} is an idiosyncratic error term.

Table 3 shows that none of the characteristics are correlated with classroom peers' characteristics.

Teacher Assignment A related concern is that certain homeroom teachers are more likely to be assigned to the students with a higher propensity to gain weight, in which case our peer effects estimates will be confounded with teacher effects. There is no evidence that teachers in South Korea are assigned to classrooms in a way that might lead to this type of bias (Lim and Meer, forthcoming). However, we examine whether observable teacher characteristics are correlated with student characteristics. Of the 25 coefficients from regressions of a student characteristic on teacher characteristics, only one is statistically significant at the 10 percent level. This provides strongly suggestive evidence that teachers are not assigned to classrooms in a systematic manner.

Specification

The basic linear-in-means regression equation we use to estimate the effects of peers' BMI is:

$$BMI_{ics}^{7th} = \alpha + \frac{1}{n_c - 1} \sum_{j \neq i, j \in c(i)} (\beta BMI_{jcs}^{7th} + X_{jcs}\gamma') + X_{ics}\delta' + HT_{cs}\theta' + \lambda_s + \varepsilon_{ics},$$

(2)

where BMI_{ics}^{7th} is the BMI for seventh grade student i, assigned to classroom c in school s. $\frac{1}{n_c-1} \sum_{j \neq i, j \in c(i)} BMI_{jcs}^{7th}$ is the average BMI of student i's peers. X_{ics} and HT_{cs} are vectors of student and homeroom teacher characteristics. Student characteristics include number of siblings, student gender, family income, whether both parents live together, whether parents own their home, and whether parents have bachelor's degrees or higher. $\frac{1}{n_c-1} \sum_{j \neq i, j \in c(i)} X_{jcs}$ is peers' average characteristics. Homeroom teacher characteristics include teacher

gender and dummies for teacher experience of less than five years, having a graduate degree, and graduating from a teachers college. λ_s are school fixed effects to account for the random assignment of students into classroom within a school; they also account for any broader unobserved impacts at the school level, such as its location. ε_{ics} is the error term.

Instrumental Variable

β is intended to capture the endogenous peer effects – that is, the impact of peers on the individual (Manski, 1993). However, this coefficient will be biased because of the reflection problem and contextual effects; namely, the student and his or her peers affect each other's BMI simultaneously, and they share the same environment (the classroom). As such, a naïve estimate will confound these effects. To address this problem, we instrument peers' BMI with peers' number of siblings; a number of studies find that the number of siblings is correlated with the child's own BMI and the likelihood of being obese (Hesketh et al., 2007; Chen and Escarce, 2010; Haugaard et al., 2013; de Oliveira Meller et al., 2015). Children with more siblings tend to have a higher level of physical activity and lower food intake (Jacoby et al., 1975; Bagley et al., 2006). Parenting practices may differ for multi-child families, with single-child families having more restrictions on outdoor play and fewer restrictions on food (Hesketh et al., 2007). Figure 1, which depicts average BMI by the number of siblings in our data, shows a clear negative relationship for children with zero to five siblings (we exclude the small number of observations with six or more siblings).

The number of peers' siblings is thus clearly relevant to peers' BMI, satisfying the first condition for an instrumental variable. The instrument must also not directly affect the outcome variable – that is, the number of siblings that peers have cannot directly impact the student's own BMI except through its impact on peer BMI. It is difficult to imagine a mechanism by which this could be the case, though we remain agnostic on the exact nature of this transmission. Peer BMI is the outcome of numerous inputs, including genetics, nutrition, and physical

9

activity. If peer BMI changes because of an increase in physical activity, that may have a different impact than if it changes because of an improvement in dietary habits.

We examine whether peers' number of siblings is correlated with student or teacher characteristics, which would provide evidence against random assignment. Table 4 shows that only the coefficient from a regression of the instrument on family income is significant at 10 percent level.

We therefore use the following as a first stage equation in a two-stage least squares frameworking, instrumenting peer BMI with the number of peer siblings. We also examine reduced-form results, in which peer number of siblings is substituted for peer BMI in the equation.

$$\frac{1}{n_c - 1} \sum_{j \neq i, j \in c(i)} BMI_{jcs}^{7th} = \gamma + \delta \frac{1}{n_c - 1} \sum_{j \neq i, j \in c(i)} S_{jcs} + \theta S_{ics} + \eta_s + \epsilon_{ics},$$

$$(3)$$

where S_{ics} and S_{jcs} indicate number of siblings for student i and $j \neq i$.

Results

Contemporaneous Effects

Table 5 presents the coefficients from estimating variations of Equation 1. Column 1 reports the naïve ordinary least squares result. The large negative effect seems to suggest that heavier peers cause a *reduction* in BMI. However, as discussed above, random assignment and balance across classrooms will mechanically generate a negative effect.* To illustrate this point clearly, we examine

---

*Suppose the sum of students' BMIs and the number of students in each classroom are $\bar{y}$ and n and the same across classrooms. We define average classroom peers' BMI for student i as $X_i = \frac{\bar{y} - y_i}{n-1}$, where y_i is own BMI. Then, a regression of own BMI on peers' BMI is the same as estimating the following equation:

$$y_i = \alpha + \beta X_i$$
$$= \alpha + \beta(\frac{\bar{y} - y_i}{n-1})$$
$$= \alpha + \beta \frac{\bar{y}}{n-1} - \beta \frac{1}{n-1} y_i$$

whether the OLS coefficient in Column 1 is more extreme than synthetic coefficients made of randomly resampled students from the same school. We take students within a school and randomly reassign them into the classrooms of the same size as the existing classrooms. Using the artificial classrooms, we estimate the same specification as in Column 1 of Table 5 10,000 times, obtaining the sample distribution of the coefficients under the null hypothesis that peers' BMI is random. Figure 2 shows that the median OLS coefficient from this exercise is -0.964; 2,631 (26.3%) coefficients in the distribution that are greater than or equal to our estimate of -0.742. We therefore conclude the observed coefficient is in line with the mechanical bias of peer effects due to random assignment.

We next turn to the reduced-form results in Column 2. This substitutes the number of peer siblings for peer BMI in the regression and provide a first pass at the impact of peer BMI. An increase of one in the average number of peer siblings (1.37 standard deviations) reduces the student's own BMI by a little over one unit (0.36 standard deviations). That is, students who have peers with more siblings have, on average, lower BMI.

In Columns 3 through 6, we show two-stage least squares (2SLS) results, which rest on the assumption that the variation in peers' BMI induced by siblings causes changes in the student's own BMI. That is, introducing more high-BMI peers will cause a student's BMI to increase. We do not suggest that BMI itself is "contagious," but rather that the habits associated with higher-BMI peers have an impact on one's own habits and therefore BMI.

Panel B shows the first-stage regression, with peer siblings exerting a significant impact on peer BMI, as expected. These results are robust to the additional of student controls, peer controls, and teacher controls in each subsequent column. Panel A shows the second stage of the results, with a one-unit increase in peers' average BMI increasing own BMI by 0.83 units. This estimate is the opposite sign of the OLS estimate in Column 1, and is precisely estimated. It does not vary much with the addition of controls in Columns 4-6, indicating

---

Because $-\beta\frac{1}{n-1}$ is always 1, $\beta = -(n-1) < 0$.

that peer BMI is uncorrelated with student, peer, and teacher characteristics and providing further evidence that random assignment holds in our data.

The effect is somewhat larger than that in Trogdon, Nonnemaker and Pais (0.52) and Mora and Gil (0.63) (Trogdon et al., 2008; Mora and Gil, 2013). Christakis and Fowler examine the effect of peers on obesity rather than BMI, finding that the chance of being obese increases by 57 percent if a friend is obese (Christakis and Fowler, 2007). Our larger effects might be due in part to the exposure of classroom peers to each other – that is, during the school day, throughout the entire year; as one would expect, peer effects are larger when the peer group is defined as those with whom an individual spends more time (Carrell et al., 2009)

Next, we examine the effect of peers' BMI on the likelihood of a student being overweight or obese. We use the definition of adolescents being obese or overweight from the 2007 WHO Reference Chart for Boys and Girls, which varies by gender and age. In Table 6, the dependent variable in Column 1 is an indicator for a student's BMI being greater than or equal to the overweight cutoff and shows that a one-unit increase in peer BMI increases the likelihood of being overweight by 10.8 percentage points, a meaningfully large impact given that, on average, 19.9 percent of 7th graders are overweight. Column 2 is the same specification for an indicator for the probability of being obese; the impact of heavier peers is positive but imprecisely estimated.

Effects Over Time

We examine the effects of seventh grade peers' BMI on eighth and ninth grade BMI to see if the effects in Table 5 fade out. We are able to isolate seventh grade peer effects from the later year's peer effects because students are again randomly assigned to a new classroom at the beginning of the new academic year; however, the structure of GEPS2012 does not allow us to re-estimate contemporaneous effects, since eighth and ninth grade peers are not fully sampled. In Table 7, we regress student's weight outcomes in eighth or ninth grades on

peers' BMI in seventh grade controlling for seventh grade school fixed effects and instrumenting for peer BMI with peer number of siblings.

The first column shows that seventh grade peers on BMI still have an impact in eighth grade, but about a third of the initial effect of about 0.8 has faded out. By ninth grade, the impact of seventh grade peers no longer has a significant impact on BMI. However, there is still a positive and significant impact on the likelihood of being overweight even in ninth grade, suggesting that certain children are particularly susceptible to peer influence on BMI (note, however, that there are no significant effects for obesity). That is, the effects are transitory for most students, but those who are at risk of being overweight are more likely to remain so. While we do not take a strong position on this issue, the relatively small impacts on BMI coupled with the fairly substantial remaining effects on the likelihood of being overweight can be reconciled by heterogeneity in treatment effects.

Falsification Test

Cohen-Cole and Fletcher (2008a) suggest that finding peer effects on implausible characteristics, such as height, indicates flawed methodology and spurious correlation. BMI is composed of both height and weight; there is no reason why peers' BMI should have a causal impact on a student's height. We therefore separate height and weight and estimate the specifications in Table 5 in Table 8. Panel A reports the estimates on height in centimeters. The OLS coefficient in Column 1 is negative, as is to be expected. There is no impact of peers' height in the reduced form and 2SLS specifications; in Column 6, a one-unit increase in average peer BMI is associated with 0.2 centimeter lower height, both statistically and economically insignificant. In Panel B, we focus on weight alone. In this case, the impact is large and precisely estimated, with a one unit increase in peer BMI causing an increase of slightly over 2 kilograms in own weight. Altogether, these findings provide more evidence that our results are not spurious.

Discussion

We examine the effects of peers' BMI on middle school student's own weight outcomes. The random assignment of peers in our data, combined with our instrumental variables approach, enables us to identify endogenous peer effects. We find that a one-unit increase in peers' average BMI increases a student's own BMI by 0.83 units and the likelihood of being overweight by 10.8 percentage points. We show that peers BMI still have an impact one year after exposure. Effects on the likelihood of being overweight last longer, suggesting that the peer effects are transitory for most students, but persistent for those who are at risk of being overweight.

Recent findings that obesity spreads through social network have been criticized for not properly handling the challenges of non-random peer groups. Our natural experiment, in which students are randomly assigned their classroom peers, circumvents this problem. We instruments peer BMI with peers' number of siblings, which is negatively correlated with peers BMI but clearly independent from own BMI, to account for contextual effects, the reflection problem, and measurement error. We show that peer height exhibits no contagion, while peer weight does, suggesting that our results are not spurious. Further, we document a mechanical negative relationship between peers' and own BMI when we reassign students into artificial classrooms, providing further evidence that our results are not driven by confounds.

Our robust results on the social network effects provide strong evidence that social factors play a large role in the obesity epidemic, suggesting the importance of approaching obesity as both a public health and a clinical problem. That is, policy intervention to reduce overweight or obesity can have multiplier effects by influencing on both targeted subjects and their peers, particularly for adolescents.

References

Bagley, Sarah, Jo Salmon, and David Crawford, "Family structure and children's television viewing and physical activity," *Medicine and Science in Sports and Exercise*, 2006, *38* (5), 910–918.

Bellizzi, Mary C. and William H. Dietz, "Workshop on childhood obesity: summary of the discussion," *The American Journal of Clinical Nutrition*, 1999, *70* (1), 173s–175s.

Biener, Adam, John Cawley, and Chad Meyerhoefer, "The Medical Care Costs of Youth Obesity: An Instrumental Variables Approach," *NBER Working Paper No. 23682*, 2017.

Burkhauser, Richard V. and John Cawley, "Beyond BMI: the value of more accurate measures of fatness and obesity in social science research," *Journal of Health Economics*, 2008, *27* (2), 519–529.

Carrell, Scott E., Mark Hoekstra, and James E. West, "Is poor fitness contagious? Evidence from randomly assigned friends," *Journal of Public Economics*, 2011, *95* (7), 657–663.

_ **, Richard L. Fullerton, and James E. West**, "Does your cohort matter? Measuring peer effects in college achievement," *Journal of Labor Economics*, 2009, *27* (3), 439–464.

Cawley, John, Euna Han, Jiyoon Kim, and Edward Norton, "Testing for Peer Effects Using Genetic Data," *NBER Working Paper No. 23719*, 2017.

Chen, Alex Y and José J Escarce, "Family structure and childhood obesity, Early Childhood Longitudinal Study – Kindergarten Cohort," *Prev Chronic Dis*, 2010, *7* (3), A50.

Christakis, Nicholas A and James H Fowler, "The spread of obesity in a large social network over 32 years," *New England Journal of Medicine*, 2007, *357* (4), 370–379.

Cohen-Cole, Ethan and Jason M. Fletcher, "Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis," *British Medical Journal*, 2008, *337*, a2533.

_ **and** _ , "Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic," *Journal of Health Economics*, 2008, *27* (5), 1382–1387.

de Oliveira Meller, Fernanda, MCF Assunção, AA Schäfer, CL de Mola, AJD Barros, DL Dahly, and FC Barros, "The influence of birth order and number of siblings on adolescent body composition: evidence from a Brazilian birth cohort study," *British Journal of Nutrition*, 2015, *114* (01), 118–125.

Epple, Dennis and Richard Romano, "Peer effects in education: a survey of the theory and evidence," *Handbook of Social Economics*, 2011, *1* (11), 1053–1163.

Fowler, James H and Nicholas A. Christakis, "Estimating peer effects on health in social networks: a response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais," *Journal of Health Economics*, 2008, *27* (5), 1400.

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo, "Peer effects in the workplace: evidence from random groupings in professional golf tournaments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 34–68.

Haugaard, Line K, Teresa A Ajslev, Esther Zimmermann, Lars Ängquist, and Thorkild IA Sørensen, "Being an only or last-born child increases later risk of obesity," *PLoS One*, 2013, *8* (2), e56357.

Hesketh, Kylie, David Crawford, Jo Salmon, Michelle Jackson, and Karen Campbell, "Associations between family circumstance and weight status of Australian children," *International Journal of Pediatric Obesity*, 2007, *2* (2), 86–96.

Jacoby, Ann, D. G. Altman, Judith Cook, W. W. Holland, and Elliottm A., "Influence of some social and environmental factors on the nutrient intake and nutritional status of schoolchildren," *British Journal of Preventive and Social Medicine*, 1975, *29* (2), 116–120.

Kang, Changhui, "Classroom peer effects and academic achievement: quasi-randomization evidence from South Korea," *Journal of Urban Economics*, 2007, *61* (3), 458–495.

Kuczmarski, Marie Fanelli, Robert J Kuczmarski, and Matthew Najjar, "Effects of age on validity of self-reported height, weight, and body mass index: findings from the Third National Health and Nutrition Examination Survey, 1988–1994," *Journal of the American Dietetic Association*, 2001, *101* (1), 28–34.

Lee, Dong-Hun, Aesun Shin, Jeongseon Kim, Keun-Young Yoo, and Joohon Sung, "Validity of self-reported height and weight in a Korean population," *Journal of epidemiology*, 2011, *21* (1), 30–36.

Lim, Jaegeum and Jonathan Meer, "The impact of teacher-student gender matches: random assignment evidence from South Korea," *Journal of Human Resources*, forthcoming.

Manski, Charles F., "Identification of endogenous social effects: the reflection problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Mora, Toni and Joan Gil, "Peer effects in adolescent BMI: evidence from Spain," *Health Economics*, 2013, *22* (5), 501–516.

OECD, *Obesity and the economics of prevention*, OECD Publishing, 2010.

_ , *Health at a glance 2015*, OECD Publishing, 2015.

Strauss, Richard S, "Comparison of measured and self-reported weight and height in a cross-sectional sample of young adolescents.," *International Journal of Obesity & Related Metabolic Disorders*, 1999, *23* (8).

Trogdon, Justin G, James Nonnemaker, and Joanne Pais, "Peer effects in adolescent overweight," *Journal of Health Economics*, 2008, *27* (5), 1388–1399.

Yakusheva, Olga, Kandice A Kapinos, and Daniel Eisenberg, "Estimating heterogeneous and hierarchical peer effects on body weight using roommate assignments as a natural experiment," *Journal of Human Resources*, 2014, *49* (1), 234–261.

Table 1: Summary Statistics

Variable	Mean	(Std. Dev.)	Min	Max	N
A. Student Characteristics					
BMI in 7th Grade	19.4	(3.0)	11.8	34.1	3,909
8th Grade	19.9	(3.1)	10.5	39.5	3,556
9th Grade	20.2	(3.1)	12.6	35.9	3,467
Weight in 7th Grade (kg)	49.7	(9.6)	26	99	3,909
8th Grade	53.8	(10.4)	30	120	3,561
9th Grade	55.8	(10.7)	32	110	3,476
Height in 7th Grade (cm)	159.9	(6.9)	128	184	3,909
8th Grade	164.1	(7.3)	136	198	3,701
9th Grade	165.7	(7.7)	142	192	3,608
Number of Siblings	1.15	(0.73)	0	17	3,909
Female Student	0.48	(0.50)	0	1	3,909
Dad with BA or more	0.43	(0.50)	0	1	3,801
Mom with BA or more	0.30	(0.46)	0	1	3,775
Both Parents	0.84	(0.36)	0	1	3,908
Own Home	0.63	(0.48)	0	1	3,845
Family Income (million KRW)	4.78	(5.01)	0	99.99	3,837
B. Peer Characteristics					
Peers' BMI in 7th Grade	19.4	(0.7)	17.1	21.8	3,909
8th Grade	19.9	(0.7)	18.2	22.8	3,556
9th Grade	20.2	(0.7)	18.6	22.5	3,467
Peers' Weight in 7th Grade (kg)	49.7	(2.2)	43.1	57.1	3,909
8th Grade	53.8	(2.5)	46.7	64.9	3,561
9th Grade	55.8	(2.6)	48.9	63.9	3,476
Peers' Height in 7th Grade (cm)	159.9	(1.5)	154.4	165.8	3,909
8th Grade	164.0	(1.8)	156.5	170.4	3,701
9th Grade	165.6	(2.0)	157.6	173.0	3,608
Peers' Number of Siblings	1.14	(0.15)	0.74	1.81	3,909
C. Homeroom Teacher Characteristics					
Female Teacher	0.82	(0.39)	0	1	3,768
Teacher over 40	0.29	(0.45)	0	1	3,768
Teacher Experience below 5 Yrs	0.21	(0.41)	0	1	3,768
Teacher's College	0.69	(0.46)	0	1	3,732
Post Graduate Teacher	0.39	(0.49)	0	1	3,768
Administrative Teacher	0.06	(0.24)	0	1	3,632

Table 2: Number of Rejections in Pearson's χ^2 Tests

Significance Level	7th Grade (367 Tests)			8th Grade (372 Tests)			9th Grade (372 Tests)		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
Female Student	0	0	0	0	0	1	0	1	3
Number of Siblings	1	1	2	0	2	4	0	1	3
Dad Education	0	0	1	0	2	3	2	4	7
Mom Education	0	0	2	1	5	8	1	2	5
Married Parents	0	2	7	1	3	4	0	4	6
Own Housing	1	3	4	0	2	6	1	2	8
Sum	2	6	16	2	14	26	4	14	32
Percent Rejected	0.5	1.6	4.4	0.5	3.8	7.0	1.1	3.8	8.6

Table 3: Regression of Own Characteristics on Peer Characteristics

	(1) FS	(2) # of Siblings	(3) Parents BA +	(4) Married Parents	(5) Owning Housing	(6) Familiy Income
Avg. Class Peers	−0.115	−0.021	−0.041	−0.046	0.034	−0.016
Characteristic	(0.138)	(0.049)	(0.032)	(0.061)	(0.031)	(0.019)
Observations	4,046	4,046	3,803	4,006	3,973	3,965
R^2	0.955	0.957	0.956	0.941	0.951	0.951

Notes: Each column represents a separate regression, including school fixed effects and the leave-out mean of school peers' characteristic, which is defined as school mean for the relevant characteristic excluding the individual's own value. We include the school peers' leave-out mean to correct the bias that is inherent in typical tests for random assignment of peers. Dependent variables in Columns 1 through 6 are an indicator for female student, number of siblings, indicators for both parents having B.A. degree or higher, for parents being married, for parents owning their home, and family income. Standard errors in parentheses are clustered at school level.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 4: Regressions of Characteristics on Peer Number of Siblings

A. Own Student Characteristics

	Female Student (1)	Dad BA + (2)	Mom BA + (3)	Both Parents (4)	Own Home (5)	Family Income (6)
Peers' # of Siblings	−0.043	−0.038	−0.000	−0.057	0.003	−1.776*
	(0.036)	(0.057)	(0.061)	(0.065)	(0.083)	(0.910)
Constant	0.579***	0.515***	0.302**	0.966***	0.622***	8.573***
	(0.076)	(0.123)	(0.130)	(0.140)	(0.178)	(1.950)
Observations	4,037	3,924	3,897	4,036	3,969	3,962

B. Homeroom Teacher Characteristics

	Female Teacher (1)	Teacher over 40 (2)	Low Experience (3)	Teacher's College (4)	Post Graduate (5)	Admin Teacher (6)
Peers' # of Siblings	−0.189	0.356	−0.205	0.720	0.485	−0.042
	(0.399)	(0.285)	(0.286)	(0.468)	(0.375)	(0.258)
Constant	1.219	−0.476	0.645	−0.855	−0.651	0.153
	(0.853)	(0.610)	(0.612)	(1.002)	(0.802)	(0.554)
Observations	3,888	3,888	3,888	3,851	3,888	3,747

Notes: Each column represents a separate regression, controlling for school fixed effects. In Panel A, dependent variables are indicators for female student, dad with bachelor's degree or higher, mom with bachelor's degree or higher, both parents living together, parents' owning home, and family income in millions of KRW (930 USD at the end of 2012). In Panel B, dependent variables are indicators for female teacher, whether the teacher is older than 40, whether the teacher has less than 5 years of experience, whether the teacher graduated from a teachers college, whether the teacher has a postgraduate degree, and whether the teacher holds an administrative position. Standard errors in parentheses are clustered at school level.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 5: Peer Effects on BMI

	OLS	RF	2SLS			
	(1)	(2)	(3)	(4)	(5)	(6)
A. OLS & 2SLS Results						
Avg. Peers' BMI	−0.742***		0.831***	0.792***	0.806***	0.882***
	(0.262)		(0.095)	(0.221)	(0.155)	(0.165)
Peer # of Siblings		−1.093**				
		(0.494)				
Own # of Siblings		−0.241***	−0.214***	−0.189***	−0.189***	−0.212***
		(0.071)	(0.075)	(0.067)	(0.068)	(0.068)
Observations	3,909	3,909	3,909	3,584	3,584	3,418
R^2	0.052	0.046	0.008	0.019	0.019	0.015
B. 1st Stage Results						
Peer # of Siblings			−1.315***	−1.275***	−1.439***	−1.499***
			(0.480)	(0.478)	(0.453)	(0.511)
Own # of Siblings			−0.033**	−0.028*	−0.035**	−0.035*
			(0.017)	(0.017)	(0.016)	(0.019)
1st Stage F-stat			7.52	7.12	10.10	8.59
School FEs	Yes	Yes	Yes	Yes	Yes	Yes
Student Ctrl				Yes	Yes	Yes
Peer Ctrl					Yes	Yes
Teacher Ctrl						Yes

Notes: Each column represents a separate regression. The dependent variable is a student's BMI and independent variable in the 2SLS model is mean peer BMI, which is instrumented with average peers' number of siblings. Student characteristics include student gender, family income, and indicators for both parents living together, father having a bachelor's degree or higher, and mother having a bachelor's degree or higher, and parents owning their home. Homeroom teacher characteristics include teacher gender and indicators for less than five years' experience, having a graduate degree, and graduating from a teachers' college. Peer characteristics are leave-out means of student characteristics. Standard errors in parentheses are clustered at school level.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 6: Peer Effects on the Probability of Being Overweight or Obese

| | Dep. Var = Indicator for being | |
	Overweight	Obese
Avg. Peers' BMI	0.108***	0.032
	(0.026)	(0.022)
Own Number of Siblings	−0.034***	−0.007*
	(0.008)	(0.004)
Observations	3,909	3,909
R^2	0.013	0.012
1st Stage F-stat	7.54	7.54

Notes: Each column represents a separate regression, controlling for school fixed effects. Dependent variables in Columns 1 and 2 are indicators for a student being overweight and being obese, respectively. 7th grade peer BMI is instrumented with 7th grade peers' number of siblings. Standard errors in parentheses are clustered at school level.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 7: Peer Effects over Time

| | BMI | | Overweight | | Obesity | |
	8th	9th	8th	9th	8th	9th
7th Grd Avg. Peers' BMI	0.514*	0.149	0.143***	0.073*	0.013	−0.013
	(0.273)	(0.294)	(0.044)	(0.042)	(0.025)	(0.022)
Own Number of Siblings	−0.197**	−0.190**	−0.020***	−0.022***	−0.009**	−0.008*
	(0.078)	(0.077)	(0.008)	(0.008)	(0.004)	(0.004)
Observations	3,556	3,467	3,556	3,467	3,556	3,467
R^2	0.021	0.043	0.001	0.034	0.018	0.036
1st Stage F-stat	7.94	7.28	7.94	7.28	7.94	7.28

Notes: Each column represents a separate regression, controlling for 7th grade school fixed effects. Dependent variables are own BMI in 8th and 9th grades, the likelihood of being overweight in 8th and 9th grades, and the likelihood of being obese in 8th and 9th grades, respectively. The independent variable is mean peer BMI in 7th grade, instrumented with average peers' number of siblings in a 2SLS model. Standard errors in parentheses are clustered at school level.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 8: Are Peer Effects Driven by Height?

	OLS	RF	2SLS			
	(1)	(2)	(3)	(4)	(5)	(6)
A. Effect on Height in cm						
Avg. Peers' BMI	−0.569**		0.176	0.080	−0.197	−0.447
	(0.271)		(0.758)	(0.897)	(0.823)	(0.893)
Peer # of Siblings		−0.232				
		(0.994)				
Number of Siblings		−0.442***	−0.436***	−0.212	−0.207	−0.205
		(0.161)	(0.159)	(0.179)	(0.179)	(0.192)
Observations	3,909	3,909	3,909	3,584	3,584	3,418
1st Stage F-stat			7.52	7.12	10.10	8.59
B. Effect on Weight in kg						
Avg. Peers' BMI	−2.216***		2.434***	2.261***	2.124***	2.228***
	(0.794)		(0.422)	(0.822)	(0.630)	(0.689)
Peer # of Siblings		−3.202**				
		(1.402)				
Number of Siblings		−0.913***	−0.833***	−0.619**	−0.617**	−0.678***
		(0.255)	(0.258)	(0.249)	(0.249)	(0.260)
Observations	3,909	3,909	3,909	3,584	3,584	3,418
1st Stage F-stat			7.52	7.12	10.10	8.59
School FEs	Yes	Yes	Yes	Yes	Yes	Yes
Student Ctrl				Yes	Yes	Yes
Peer Ctrl					Yes	Yes
Teacher Ctrl						Yes

Notes: Each column represents a separate regression. Dependent variables are a student's height and weight in Panel A and B, respectively. Independent variables in the 2SLS model are, respectively, mean peers' height and weight in Panel A and B, which are instrumented with average peers' number of siblings. Standard errors in parentheses are clustered at school level.
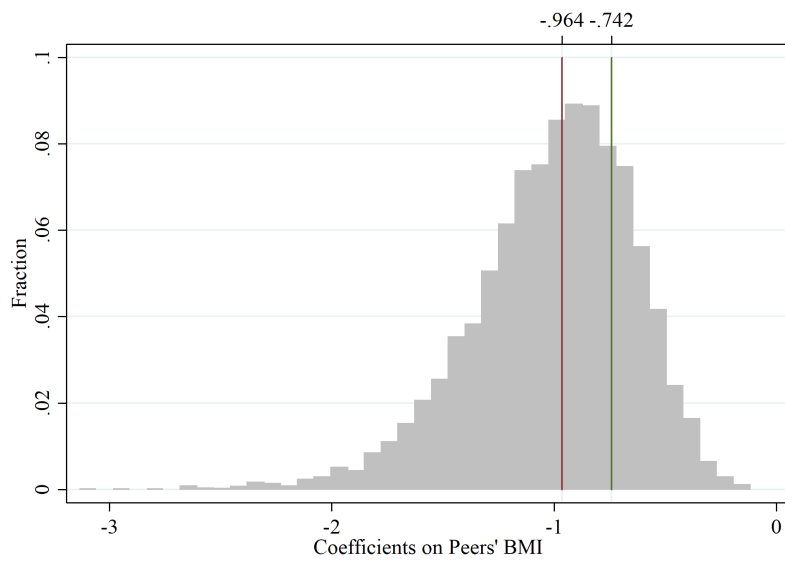
* $p < .10$, ** $p < .05$, *** $p < .01$

Figure 1: Average BMI by Number of Siblings
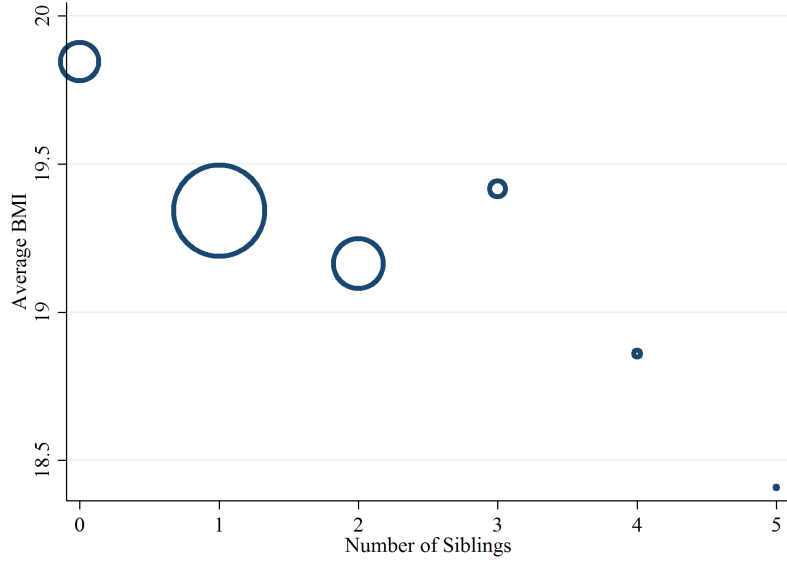
Notes: Marker size is proportional to the number of observations. We drop four points with only one observation.

Figure 2: Distribution of Coefficients under Random Assignment

Notes: We randomly resampled students from the same school and regressed own BMI on classroom peer's BMI 10,000 times. The median coefficient is -0.964 and the OLS coefficient from our data is -0.742.