I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12763

Not Only in My Genes: The Effects of Peers' Genotype on Obesity

Giorgio Brunello
Anna Sanz-de-Galdeano
Anastasia Terskaya

NOVEMBER 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12763

Not Only in My Genes: The Effects of Peers' Genotype on Obesity

Giorgio Brunello
University of Padova and IZA

Anna Sanz-de-Galdeano
University of Alicante, CRES-UPF and IZA

Anastasia Terskaya
University of Alicante

NOVEMBER 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Not Only in My Genes: The Effects of Peers' Genotype on Obesity*

We use data from three waves of Add Health to study the short- and long-run effects of high school peers' genetic predisposition to high BMI – measured by grade-mates' average BMI polygenic scores – on adolescent and adult obesity in the U.S. We find that, in the short-run, a one standard deviation increase in peers' average BMI polygenic scores raises the probability of obesity for females by 2.8 percentage points, about half the size of the effect induced by a one standard deviation increase in one's own polygenic score. No significant effect is found for males. In the long-run, however, the social-genetic effect fades away, while the effect of one's own genetic risk for BMI increases substantially. We suggest that mechanisms explaining the short-run effect for females include changes in nutrition habits and a distorted perception of body size.

JEL Classification: D62, I1, I12

Keywords: obesity, peer effect, BMI polygenic scores, Add Health

Corresponding author:
Anna Sanz-de-Galdeano
Universidad de Alicante
Carretera de San Vicente s/n
03080 San Vicente
Alicante
Spain

E-mail: anna.sdg@ua.es

* Sanz-de-Galdeano acknowledges financial support from the Spanish Ministry of Economy and Competitiveness Grant ECO2017-87069-P. We thank the participants at the 2019 Russell Sage Foundation Summer Institute in Social-Science Genomics and the audience at seminars in Alicante and Padova for comments and suggestions.

1 Introduction

In 2015, across the OECD, 19.5% of the adult population was obese (OECD, 2017). According to the WHO, [1] the prevalence of overweight and obesity among children and adolescents aged 5-19 has risen dramatically from just 4% in 1975 to over 18% in 2016. This rise has affected in a similar fashion both boys and girls. The observed trend is a source of concern because obese children and adolescents are around five times more likely to be obese in adulthood than normal - weight children and adolescents.[2] In addition, child and adolescent overweight and obesity have been found to significantly increase the risk of premature mortality and to be associated with significantly increased risk of later cardio-metabolic morbidity in adult life (Reilly and Kelly, 2011).

As argued by Cawley (2004a, 2011), if individuals were perfectly rational and their decisions about food and weight imposed no costs on others in society, if information about the consequences of obesity were accurate and readily available and there were no externalities, there would be no market failure and no reason for government intervention on efficiency grounds. When one or more of these assumptions are violated, however, the socially optimal level of obesity can differ from the one that prevails when individuals operate in free markets. For instance, externalities can occur when individuals with high or low genetic propensity to particularly high or low BMI fail to recognize that their weight affects the weight of those with whom they interact.[3]

Cawley et al. (2017) lists five economic factors explaining childhood obesity: food prices, agricultural policies, income, maternal employment and technology. Genes also matter, as small genetic differences at birth can affect both obesity risk, health and human capital formation[4] and lead to substantial inequality in long-term social and economic outcomes (including productivity, wages and skill formation).[5]

In this paper, we investigate whether adolescent and adult obesity are affected not only by individuals' own genes, but also by the genes of peers the adolescents interact with in the same grade at school. These genes can operate both via peers' obesity and BMI and by influencing other peers' traits and behaviours (e.g. smoking) which can impact on individual behaviour.

Our research contributes to a growing literature that exploits the increased availability of genetic data to dig deeper into questions that are relevant for social scientists (Beauchamp et al., 2011; Benjamin et al., 2012; Conley and Fletcher, 2018a) and to inquire whether the ge-

[1] https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

[2] According to Simmonds et al. (2016), around 80% of obese adolescents will still be obese in adulthood. However, 70% of obese adults were not obese in childhood or adolescence.

[3] In the model by Cutler et al. (2003), individuals choose the consumption of goods affecting weight and other consumption goods to maximize utility, taking into account that higher weight reduces utility. In this setup, peers' weight can influence consumption and weight by affecting preferences and the technology that produces weight.

[4] See for instance Yang et al. (2007), Speliotes et al. (2010), Sandholt et al. (2012), Jou (2014), Locke et al. (2015), and Biroli (2015).

[5] See for instance Cawley (2004b), Brunello and d'Hombres (2007), Brunello et al. (2008), Cawley (2015), and Böckerman et al. (2019).

netic makeup of individuals around us impacts our own outcomes (Domingue and Belsky, 2017; Sotoudeh et al., 2019; Cawley et al., 2017).[6] This literature has shown that smoking, obesity, and other health indicators are shaped by the interaction between genes, environmental factors (Bierut et al., 2018; Barcellos et al., 2018), and investment choices over the life-cycle (Biroli, 2015).[7]

Previous research has also produced evidence of "genetic nurturing effects" within the family (Kong et al., 2018): children's education is affected not only directly by the parental and maternal genetic variants they inherit, but also indirectly by the genes affecting the behaviour of individuals they interact with (parents, relatives, siblings). In particular, (Cawley et al., 2017) have shown that the probability of being obese is higher for those having a sibling with a high genetic predisposition to obesity, even after controlling for one's own genetic predisposition to obesity.

Our research is motivated by the scarce evidence to date on meta-genomic effects outside the family. Two important contributions in this field are Domingue et al. (2018), who document the presence of genetic similarities among school-based friends and schoolmates, and Sotoudeh et al. (2019), who show that a significant relationship exists between the genetic predisposition to smoking of an adolescent's grade-mates, classmates and school friends and the adolescent's smoking behaviour.

While these contributions focus on short-run effects, we investigate whether social-genetic effects on obesity at school last over time or fade away as adolescents grow into adulthood. We compare the impact of high school peers' genetic predisposition to high BMI —measured with average BMI polygenic scores in the school and grade— on individual obesity both close to the time when social interactions occur (the short-run) and more than 10 years after grade completion (the long-run).[8]

The estimation of the causal effect of social interactions is challenging because of self-selection into peer groups, omitted common environmental factors, and the reflection problem (Manski, 1993). The fact that genes are determined before group formation solves both the common environment and the reflection problems. We address non-random assignment of individuals into peer groups by relying on the observation that the assignment of students to a grade within a school is as good as random, and by focusing on the genetic markers of grade-mates rather than on those of schoolmates or friends. We assess the validity of this identification strategy by showing that the actual within-school variation in average grade-mates' genetic scores is not statistically different from the variation obtained using Monte-

[6] These effects have also been referred to as "indirect" genetic effects in the literature.

[7] Other areas of investigation include the relationship between individuals' genetic predisposition to educational attainment and labour market outcomes (Papageorge and Thom, 2019), wealth inequality (Barth et al., 2019), human capital accumulation (Domingue et al., 2015; Papageorge and Thom, 2019; Ronda et al., 2019), social class mobility (Belsky et al., 2018), parental investment decisions (Breinholt and Conley, 2019; Sanz-de Galdeano and Terskaya, 2019; Wertz et al., 2019), and health (Bolyard and Savelyev, 2019).

[8] Polygenic scores are scores that predict genetic predisposing to a particular trait (e.g. BMI or educational attainment). These scores are based on genome-wide association studies (GWAS), which use information on the whole genome. As a result, the BMI polygenic score is a weighted sum of alleles significantly associated with higher BMI.

Carlo simulations.

We find that, in the short-run, grade-mates' BMI polygenic scores have a positive effect on female adolescents' obesity but no effect on male obesity.[9] Using our preferred specification, we estimate that a one standard deviation increase in these scores raises the probability of obesity for females by 2.8 percentage points, about half the size of the impact of a one standard deviation increase in their own polygenic scores. In the long-run, however, peers' BMI polygenic scores have no effect on adult obesity, neither for males nor for females.

We also show that the short-run effect of peer's BMI polygenic scores on female obesity is not linear and mostly driven by the sub-group of peers with a strong genetic predisposition to high BMI. We discuss potential mechanisms and argue that nutrition habits and a distorted perception of body size may help explain our short-run results for females.

In addition, we estimate unconditional quantile regressions to explore how social-genetic effects vary along the distribution of individual BMI. We find that, while females with high BMI are affected by their grade-mates' BMI PGS, females whose BMI is around the median or below are unaffected. The literature shows evidence supporting both genetic (Domingue et al., 2018) and weight-based homophily (Crosnoe et al., 2008; De La Haye et al., 2011; Schaefer and Simpkins, 2014). When both are present, individuals self-select into groups formed by other individuals who are similar in terms of both their BMI PGS and their BMI. In these circumstances, our results suggest that randomly assigning individuals to groups (i.e., classes) may reduce obesity prevalence among high BMI females without affecting those with low BMI.

Based on their evidence, showing that sibling peer effects influence obesity, Cawley et al. (2017) have argued that programs that prevent weight gain or facilitate weight loss could exhibit social multiplier effects by benefitting not only participants but also their peers, and that this might lead to under-estimating their cost-effectiveness.[10] But are these effects important? Using our estimates to infer the impact of peers' obesity on individual obesity, we find that a ten percentage points increase in the former raises the probability of female obesity by 3.2 percentage points in the short-run but has no significant effect in the long-run. This evidence suggests that social multiplier effects in schools are likely to be short-lived.

The remainder of the paper is organized as follows. Section 2 introduces the data, Section 3 discusses our empirical approach, Section 4 presents the results, and Section 5 presents the robustness checks . Conclusions follow.

[9] We conduct our analysis separately for females and males because social effects may vary significantly by gender, as it was found in Trogdon et al. (2008); Yakusheva et al. (2014); Loh and Li (2013), among others for BMI related outcomes, and in Argys and Rees (2008); Fletcher et al. (2013); Black et al. (2013); Rodríguez-Planas et al. (2018), among others for other outcomes, including academic achievement, labour market outcomes, and risky behaviours.

[10] Analogous arguments regarding social multiplier effects have been made in related contexts such as physical fitness (Carrell et al., 2011).

2 Data

2.1 The Add Health Dataset

The National Longitudinal Study of Adolescent to Adult Health (Add Health hereafter) is a nationally representative longitudinal survey of U.S. 7th to 12th graders during the school year 1994/95 drawn from a stratified sample of 80 high schools and 52 middle schools.[11] Add Health Wave I data collection took place between September 1994 and December 1995 (age range: 12-20 years; mean age 15.7), and included both an in-school questionnaire and an in-home interview. The in-school questionnaire was administered to all the students attending the participating schools on the interview day and gathered information on school, social and demographic characteristics of students, including their parental background. Within each school and grade, a random sample of approximately 17 males and 17 females were selected to complete a longer and more detailed in-home questionnaire.[12] Additionally, a parent, preferably the resident mother of each adolescent selected for the in-home sample, was also asked to complete a questionnaire in Wave I. The in-home sample of students were subsequently re-interviewed one year later (in 1996, Wave II, age range: 13-21; mean age 16.2), as well as in 2001/02 (Wave III, age range: 18-27; mean age 22), and in 2008 (Wave IV, age range: 24-33; mean age 28.5).

We use data from Waves I, II and IV of Add Health. We obtain baseline characteristics of students, their families, and the students' grade-mates from Wave I and objective measures of height and weight —used to compute the body mass index (BMI) and a binary variable equal to one for obesity and to zero otherwise— from Waves II and IV.[13] We also use Wave II to explore mechanisms.

Our working sample is obtained by applying four selection filters. First, since DNA collection for the large sample was performed only in Wave IV,[14] we select from a sample of 18,456 Wave I respondents with non-missing information on age, race, grade, school and sample weights the 14,480 individuals who were interviewed again in Wave IV.[15] Second, we consider only individuals with valid genetic data. Add Health obtained saliva samples from 96% of Wave IV respondents, 80% of whom consented to long term archiving and were therefore eligible for genome-wide genotyping. Applying quality control procedures after genotyping was performed leaves us with 8,409 individuals with valid genetic information.[16] Third, since our identification relies on the variation in grade-mates' genetic makeup across grades within

[11] Schools were stratified by region, urbanicity, school type (public, private, parochial), ethnic mix, and size.

[12] A Computer-Assisted Personal Interview (CAPI)/Audio Computer-Assisted Self Interview (ACASI) was administered to adolescents, and sections with more sensitive questions were asked in the self-administered portion of the interview.

[13] Wave I data includes height and weight self-reports that can also be used to construct BMI. In what follows we rely on objective measurements, but results based on Wave I data are in line with those presented in the paper. They are available upon request from the authors.

[14] At Wave III of Add Health, DNA was also collected but only from a subsample of full siblings and twins.

[15] The data on all other baseline covariates are from Wave I.

[16] Genome-wide data collection and quality control protocols are described in the Add Health documentation available at https://www.cpc.unc.edu/projects/addhealth/documentation/guides/PGS_AH1_UserGuide.pdf.

schools and requires information on pupils attending at least two different grades within the same school, we exclude both 2,018 individuals who attended grades with less than 15 grade-mates with available genetic data, and 361 individuals from schools in which only one grade satisfied this sample selection criterion. This leaves us with observations for 6,030 individuals. Finally, we further restrict our sample to individuals with valid information on objective indicators of BMI measured at Waves II and IV. Since not all Wave I participants could be interviewed in all subsequent waves, the final sample for Waves II and IV consists of 4,450 and 5,957 observations with non-missing values of our dependent variables.[17] Table B.1 in Appendix B shows descriptive statistics for the baseline covariates measured in Wave I.

2.2 BMI Polygenic Scores

The Add Health genetic dataset includes BMI polygenic scores (BMI PGS), a summary indicator of individuals' genetic predisposition to have a high BMI. Polygenic scores capture the genetic propensity of an individual to a particular phenotype, trait or outcome (e.g., BMI, educational attainment, etc.). This propensity —or gene-phenotype association— is identified by geneticists using genome-wide association studies (GWAS). GWAS use a data mining approach to scan the entire genome and identify the single nucleotide polymorphisms (SNPs)[18] that are associated with a particular outcome.[19] More specifically, a GWAS typically regresses an outcome of interest (such as BMI) on hundreds of thousands of individual SNPs and identifies genome-wide significant associations using conservative thresholds for statistical significance.

The polygenic scores for BMI included in Add Health rely on the results of a recent GWAS conducted by Locke et al. (2015) on a sample of 332,206 individuals. In this study, Locke et al. (2015) identified 97 BMI-associated SNPs, which are significantly associated with BMI even after adjusting p-value thresholds downward ($P - value < 5 \times 10^{-8}$) to avoid the problems associated with multiple hypothesis testing. The BMI PGS are constructed for each individual by selecting these SNPs and by computing the following indicator:

$$BMIPGS_i = \sum_{j=1}^{k} \beta_j SNP_{ij} \tag{1}$$

, where β_j is the weight estimated by Locke et al. (2015) for each SNP significantly associated with BMI.

In our working sample, BMI polygenic scores account for 5.5% (Wave II) and 5.2% (Wave IV) of the total variation in BMI. Figure 1 plots the (kernel-smoothed) densities of Add Health respondents' BMI PGS, separately by gender. The distributions are approximately normal, and do not vary significantly by gender. The estimated correlation between grade-mates'

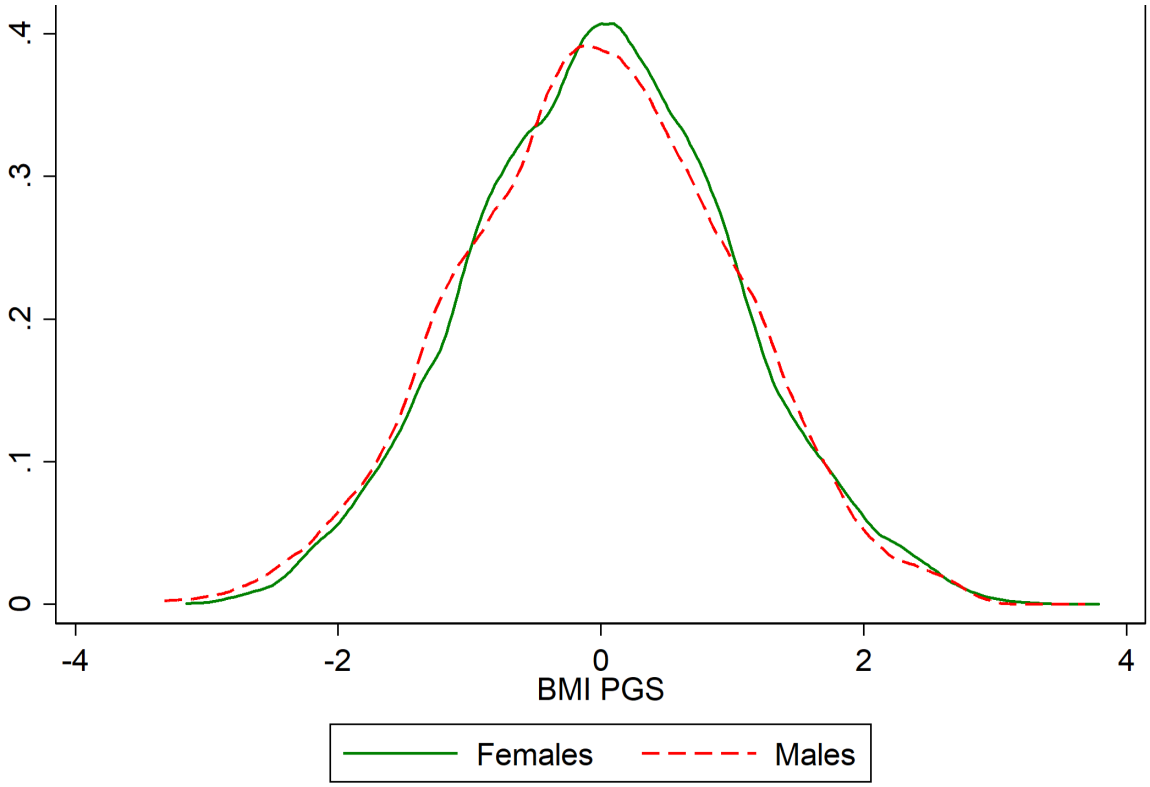
[17] Note that some Wave I respondents were not interviewed in Wave II but were interviewed at later waves. For detail see the Add Health documentation available at https://www.cpc.unc.edu/projects/addhealth/documentation/guides/DesignPaperWave_IIV.pdf.

[18] SNPs are locations in the genome where there is variation across individuals in the single nucleotide pair that is present.

[19] See https://www.yourgenome.org/stories/genome-wide-association-studiesfordetails.

average BMI PGS and grade-mates' average BMI in Wave II is 0.28, larger than the correlation between individual BMI PGS and BMI (0.20), which suggests that social multiplier effects may be present.

Figure 1: BMI Polygenic Scores (Normalized). Kernel Density Estimates



Note: Number of observations: 6030. This figure displays the kernel-smoothed densities of Add Health respondents' BMI PGS by gender.

Since the association between SNPs and outcomes as identified by GWAS studies could reflect differences in characteristics across ethnic ancestry groups rather than biological processes at the individual level, we follow standard practice in the genetics literature and account for population stratification by including as covariates in all our regressions the principal components of the full matrix of SNP data (Price et al., 2006; Benjamin et al., 2012).

An important feature of BMI PGS is that it can affect multiple traits (pleiotropy) (Cawley et al., 2011; Conley and Fletcher, 2018b). The BMI-associated SNPs identified by Locke et al. (2015) are linked, as one would expect, with other cardio-metabolic phenotypes and with outcomes such as schizophrenia, smoking behaviour, irritable bowels syndrome, and Alzheimer's disease (see Locke et al. 2015, page 4). In addition, Locke et al. (2015) find that BMI-associated SNPs overlap with genes and pathways implicated in neuro-development. Therefore, we do not claim that grademates' BMI PGS affects individual obesity only via grade-mates' BMI because the former may also affect the latter via grade-mates' smoking or other behaviours. Pleiotropy also rises concerns about the use of BMI PGS as an instrument for peers' BMI (Cawley et al., 2011, 2017).

Table 1: Summary of Outcomes

	Males and Females			Female			Males		
	mean	st.dev.	count	mean	st.dev.	count	mean	st.dev.	count
BMI. WII	23.250	5.333	4450	23.221	5.453	2343	23.280	5.205	2107
Obese. WII	0.138	0.345	4450	0.126	0.332	2343	0.151	0.358	2107
Overweight. WII	0.275	0.446	4450	0.277	0.447	2343	0.273	0.446	2107
Underweight. WII	0.038	0.191	4450	0.034	0.181	2343	0.042	0.200	2107
BMI WIV	29.191	7.817	5957	29.499	8.447	3139	28.873	7.096	2818
Obese. WIV	0.375	0.484	5957	0.394	0.489	3139	0.355	0.479	2818
Overweight. WIV	0.664	0.472	5957	0.644	0.479	3139	0.684	0.465	2818
Underweight. WIV	0.014	0.119	5957	0.020	0.141	3139	0.008	0.090	2818
BMI PGS (Normalized)	0.000	1.000	6030	0.025	0.987	3176	-0.026	1.013	2854

Note: The means are weighted using Wave I weights.

2.3 Outcome measures

We are interested in the effects of grade-mates' BMI PGS on obesity both in adolescence (average age: 16.2 years) and in adulthood (average age: 28.5 years), and in several mediating variables that may shed light on the mechanisms through which these effects operate.

Individual BMI is computed for each respondent in Waves II and IV using objective measures of weight and height and the standard formula: mass (kg) divided by squared height (m) (kg/m2).[20] Although BMI is computed in the same way for children, teens and adults, its interpretation may vary because weight, height, and their relation to body fatness change with age. For Wave II, when most respondents were still teenagers, we follow the guidelines set by the U.S. Centre for Disease Control and Prevention (Kuczmarski, 2002) and construct the following categories: underweight (BMI below the 5th percentile of the distribution), normal or healthy weight (BMI between the 5th and the 85th percentile), overweight (BMI between the 85th and the 95th percentile), and obese (BMI equal to or greater than the 95th percentile).[21] For Wave IV, when respondents were aged 24 to 33, we use instead the standard categories: underweight (BMI below 18.5), normal or healthy weight (BMI in the range 18.5-24.9), overweight (BMI in the range 25-29.9) and obese (BMI in the range 30 and above). Summary statistics for BMI and its categories by wave are reported in Table 1.

[20]Short-term test-retest reliabilities of Add Health in-home measures of height and weight (as well as derived BMI) have been shown to be uniformly excellent and to be comparable with clinic-based estimates from the National Health and Nutrition Examination Survey (NHANES). See Hussey et al. (2015), available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4487879/.

[21]We use information on BMI percentiles by sex and age in the US from 2000 CDC growth charts (the data is publicly available at https://www.cdc.gov/growthcharts/percentile_data_files.htm). For the few Wave II respondents older than 20 we use the definition of obesity applied to adults (BMI greater than or equal to 30)

3 Empirical Strategy

We consider students in grades 7th to 12th and their grade-mates (from Wave I) to evaluate the short-run and long-run effects of peers' average BMI PGS on the individual probability of being obese. Short and long-run effects are measured when individuals are on average 0.5 (Wave II) and 12.7 years (Wave IV) older than in Wave I.

There are three well known challenges to the estimation of peer effects or endogenous social effects.[22] First, peers are not randomly selected, and individuals may be more prone to forming ties with other individuals who are similar to them in terms of both observable characteristics (such as being overweight or obese) and unobserved factors (such as the individual genetic propensity to BMI). In support of this, Domingue et al. (2018) have provided suggestive evidence of genetic homophily based on Add Health, showing that friends and schoolmates are more genetically similar to one another than randomly selected peers.[23]

While some authors have dealt with non-random selection by relying on quasi-natural experiments generating randomly assigned peers,[24] others have exploited the fact that the variation in peer composition across grades within a given school may be as good as random because assignment to a grade is mostly determined by age. We use the latter approach, which was initially proposed by Hoxby (2000) and has been extensively used to study both educational and health outcomes.[25]

Second, peers are exposed to common influences and/or environments — e.g., the presence of fast food restaurants, vending machines or gyms near their school or neighbourhoods — which may lead to similar behaviours and outcomes. These contextual effects may confound peer effects, and might be an issue even when group formation is random.[26] This problem is typically addressed by controlling for school and grade fixed effects, and by considering background indicators that are determined before group formation, as we do in our analysis.

Third, estimates may be affected by the reflection problem (individual behaviour affects peers' behaviour) (Manski, 1993). While some studies have focused on peer variables that cannot be affected by the individual,[27] others have used a combination of fixed effects and/or instrumental variables.[28] This problem is not an issue in our study because individual genes are fixed at conception and grade-mates' genetic predisposition to BMI cannot be shaped by

[22] See Fletcher (2011) for an extensive review and discussion of such challenges in the context of BMI and obesity.

[23] In particular, they provide evidence of similarity for friends and schoolmates across the entire genome as well as in terms of polygenic scores for three specific phenotypes: height, BMI and educational attainment.

[24] See for instance Yakusheva et al. (2014); Guryan et al. (2009), and Carrell et al. (2011).

[25] For instance, Arduini et al. (2019) have recently used this approach to study the relationship between the BMI of grade-mates and the onset of eating disorders during adolescence. See also Cornelissen et al. (2017) for an extension of the Hoxby approach to the workplace context.

[26] See Cohen-Cole and Fletcher (2008) for a deeper discussion of the importance of environmental factors in the obesity epidemic.

[27] See for instance Hoxby (2000), Lavy and Schlosser (2011), Bifulco et al. (2011), Bertoni et al. (2017), Rodríguez-Planas et al. (2018), and Cools et al. (2019).

[28] See for instance Trogdon et al. (2008), Mora and Gil (2013), Sotoudeh et al. (2017), Asirvatham et al. (2018), and Lim and Meer (2018).

individual behaviour.

Our baseline empirical specification is:

$$Y_{isg,w} = \beta_0 + \beta_1 \overline{PGS}_{-isg,I} + \beta_2 PGS_{isg} + X'_{isg,I}\alpha + \rho_s + \delta_g + G'_{gs,I}\phi + \epsilon_{igs,w} \quad (2)$$

where $Y_{isg,w}$ is the outcome of interest in Wave II or IV for individual i attending school s and grade g (from 7th to 12th) in Wave I; $\overline{PGS}_{-isg,I}$ denotes the average BMI polygenic score of students (excluding individual i) attending the same grade and school of individual i in Wave I; PGS_{isg} denotes own BMI PGS measures in Wave IV, that are constant over the life time; $X_{isg,I}$ is a vector of student-specific characteristics in Wave I; $G_{gs,I}$ is a vector of grade-specific characteristics in school s and δ_g and ρ_s are grade and school fixed effects.

Our focus on grade-mates as peers is justified both by the amount of time that teenagers spend at school and by the large social psychology literature stressing the importance of school interactions for adolescents.[29] Our data also indicate that around 50% of the friends nominated by both males and females in our sample are school grade-mates.

Equation (2) is a reduced-form model. Grade-mates' BMI polygenic scores in this model can affect individual outcomes not only via grade-mates' BMI but also by influencing other group behaviours (for instance smoking) and traits. In addition, peers' characteristics in Wave I can affect outcomes in Waves II or IV by having an impact on the interactions occurring between waves. In any case, evidence that grade-mates' BMI PGS affect individual obesity can be interpreted as supportive of the importance of social-genetic effects for individual behaviour, regardless of the mechanisms through which they operate.[30]

We include in vector $X_{isg,I}$ the following controls, that we compute using Wave I data: i) age, age squared, gender, race indicators, and, to account for population stratification, the 10 principal components of the full matrix of genetic data (at the individual level); ii) family income, parental education, socio-economic status, parental obesity, number of siblings, and an indicator of whether both parents live in the household; iii) grade by school level indicators, such as average age, the share of females and blacks, grade size, and average parental education; iv) average schoolmates' BMI PGS, computed at the school level by excluding the individual, to account for the mechanical correlation between own and mean grade-mates' BMI PGS.[31]

[29] See for instance Dijkstra et al. (2008) and the references therein.

[30] An analogous argument has been made by (Cawley et al., 2017) when using genetic data to study BMI peer effects among full siblings

[31] As noted by Guryan et al. (2009), since individuals cannot be their own peers, the random assignment of peers generates a negative correlation between own and peers' pre-determined characteristics (grade-mates' average BMI PGS in our case). This problem is typical of peer effects studies, and the bias induced by this mechanical correlation is decreasing in the size of the population (high school students in our application) from which peers are drawn. We follow Guryan et al. (2009) and address this problem by including the school mean of BMI PGS (net of individual's own BMI PGS) as a covariate.

Table 2: Standard deviation of grade-mates' BMI PGS

	Males and Females	Females	Males
Grademates BMI PGS (Normalized)	1.000	0.999	1.002
Grade-mates' BMI PGS residuals after removing school and grade fixed effects (FE)	0.761	0.754	0.768
Grade-mates' BMI PGS residuals after removing school FE, grade FE, and school-grade controls	0.750	0.744	0.756
No. Observations:	6030	3176	2854

Note: Estimates are weighted using Wave I weights

3.1 Assessing the validity of the identification strategy

Our identification strategy relies on the presence of sufficient variation in grade-mates' genetic makeup across adjacent grades within schools, and on the credibility of the assumption that such variation is as good as random. We show in Table 2 that about 75% of the standard deviation in grade-mates' BMI polygenic scores remains after removing school and grade fixed effects (Table 2, Row 2), and school-grade level controls as well (Table 2, Row 3), suggesting that sufficient variation remains in the data to estimate the effects of interest using model (2).

We next examine the plausibility of the assumption that within-school across-grades variation in grade-mates' average BMI PGS is as good as random. This assumption, would be violated if, for instance, there was selection into grades within a school based on grade-mates' BMI PGS, which seems unlikely because grade placement generally depends on students' birth date.

We provide two pieces of evidence supporting this assumption. First, we run balancing tests to verify whether grade-mates' average BMI PGS is unrelated to individual and family background characteristics, conditional on school and grade fixed effects and on school-grade controls. Table 3 reports the estimated coefficients of the regressions of 21 covariates on grade-mates' average BMI PGS. Only one of these coefficients is statistically significant in the pooled sample of males and females (at the 10% level), in the sample of males (at the 10% level), and in the sample of females (at the 5% level), which is less than what one would obtain by chance, indicating that grade-mates' average BMI PGS is not systematically related to observed individual and family background characteristics. We also show that own BMI PGS and grademates' average BMI PGS are uncorrelated once we control for school and grade fixed effects and for school-grade controls, suggesting that there is no selection into grades within a school based on genetic homophily (Domingue et al. (2018) shows that genetic homophily affects selection of friends and selection into school).[32]

Second, using Monte-Carlo simulations as in Lavy and Schlosser (2011) and Rodríguez-

[32]In Table B.2 in Appendix B we show that the estimated effect of grade-mates' average BMI PGS on obesity is not altered when omitting individuals' own BMI PGS, indicating that, conditional on our set of controls, there is no genetic homophily, and there is no mechanical correlation between individuals' own BMI PGS and their grade-mates' average BMI PGS.

Table 3: Balancing tests

	Males and Females		Female		Male	
	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
BMI PGS (Normalized)	-0.003	0.011	-0.012	0.015	0.006	0.014
Age	-0.006	0.012	-0.003	0.018	-0.011	0.014
Female	-0.002	0.010	0.000	0.000	0.000	0.000
Black	-0.004	0.004	-0.009	0.008	0.003	0.005
White	-0.004	0.006	0.005	0.009	-0.016*	0.008
Family income	0.006	0.008	0.004	0.013	0.007	0.008
Parents are college educated	0.004	0.007	0.001	0.012	0.004	0.010
SES index (Normalized)	0.020	0.016	0.037	0.025	-0.002	0.026
At least one parent is obese	0.005	0.007	0.007	0.012	0.002	0.012
Number of siblings	-0.010	0.026	0.022	0.043	-0.032	0.044
Both parents live in household	-0.014	0.010	-0.008	0.014	-0.020	0.018
PC 1	-0.000	0.000	-0.000	0.000	-0.000	0.000
PC 2	0.000	0.000	0.000	0.000	-0.000	0.000
PC 3	-0.000	0.000	-0.001	0.000	-0.000	0.000
PC 4	-0.000	0.000	-0.000	0.001	-0.000	0.000
PC 5	-0.001*	0.000	-0.001	0.000	-0.001	0.001
PC 6	0.000	0.000	0.000	0.001	0.000	0.000
PC 7	0.000	0.000	0.000	0.001	0.000	0.000
PC 8	-0.000	0.000	0.000	0.001	-0.000	0.000
PC 9	-0.000	0.000	0.001**	0.000	-0.001	0.001
PC 10	-0.000	0.000	0.000	0.000	-0.000	0.001
No. Observations:	6030		3176		2854	

Note: The figures in each row are coefficients from regressions that include, in addition to the average BMI polygenic score of grademates, grade fixed effects, school fixed effects, and school-grade specific controls. All non-genetic variables are measured using Add Health's Wave I. All genetic variables are measured using Add Health's Wave IV. PC1-PC10 denote the first 10 principal components of the full genetic data. Standard errors are clustered at the school level. Estimates are weighted using Wave I weights. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Planas et al. (2018), we investigate whether the within school and across grades variation in average BMI PGS is similar to the variation that would result if individual BMI PGS was randomly assigned within schools. For each student in each school, we randomly draw a BMI PGS value using a normal distribution with population mean and standard deviation equal to the actual school-specific BMI PGS mean and standard deviation. We then compute school-grade level averages of simulated BMI PGS and within school standard deviations of these averages, repeat this procedure 1000 times and compute the 95% confidence interval for the simulated within-school standard deviations. We find that for 96.5% of our schools the actual standard deviation falls within the 95% confidence interval generated by simulated data, suggesting that the within-school variation in grade-mates' average BMI PGS is similar

to the one that would be obtained with random allocation.[33]

4 Results

4.1 Short-run effects on obesity and BMI

Table 4 presents the estimated effects of own and grade-mates' BMI PGS —measured in Wave I— on the probability of being obese in Wave II, less than one year later, for all (Panel A), female (Panel B), and male (Panel C) teenagers. The estimates of β_2 —the effect of grade-mates' average BMI PGS on obesity— and β_1 —the effect of own BMI PGS— are displayed in the first and second row of each panel, respectively.

The table also shows how our results vary when covariates are added sequentially: in Column 1 we only include school and grade fixed effects as well as the individual characteristics summarized in Table B.1 of Appendix B; in Column 2 we add family level controls; in Column 3 we further incorporate the 10 principal components of the full matrix of genetic data; Column 4 displays the results of our preferred specification, which also includes the school-grade characteristics summarized in Table B.1 of Appendix B, as well as average schoolmates' BMI PGS (excluding each individual's own BMI PGS value).

We find that the short-run effect of the own BMI PGS on obesity is positive and statistically significant in all specifications and for all groups. This effect should be interpreted as the overall effect of the individual's genetic predisposition to high BMI, which partly reflects parental genes (as our genes are inherited from our parents) and works through both a pure and direct biological mechanism and by triggering environmental responses (for instance by parents).[34]

In the full sample, the short-run effect of grade-mates' BMI PGS is also positive but smaller than the own effect. Consistent with the assumption that across-grades variation in average BMI PGS is as good as random, the point estimates are remarkably stable across specifications. The estimated social-genetic effect on obesity is statistically significant at the 10 percent level of confidence in Columns 1-3 but not statistically significant in our preferred specification (Column 4).

While the effect of grade-mates' BMI PGS is not statistically significant and fairly close to zero for males, it is statistically significant at the 1% level of confidence in all specifications for females. The magnitude of the estimated effect is also much larger for females than for males. For the former, we find that a one standard deviation increase in grade-mates' average BMI PGS increases the probability of being obese by 2.8 percentage points, a 22.2 percent increase with respect to the average probability of obesity (12.6 percent of females in our Wave II sample were obese).

[33] We compute within-school standard deviations using residuals from a regression of grade-mates' average BMI PGS on school fixed effects, grade fixed effects, and school-grade controls.

[34] As argued by Jencks (1980), genetic and environmental variation are not mutually exclusive, as genetic variation often causes environmental variation. In line with this idea, Breinholt and Conley (2019), Sanz-de Galdeano and Terskaya (2019), and Wertz et al. (2019) have recently studied the relationship between individuals' genetic predisposition to educational attainment and parental investment decisions.

Table 4: The short-run (Wave II) effect of grade-mates' and own BMI PGS on adolescents' obesity

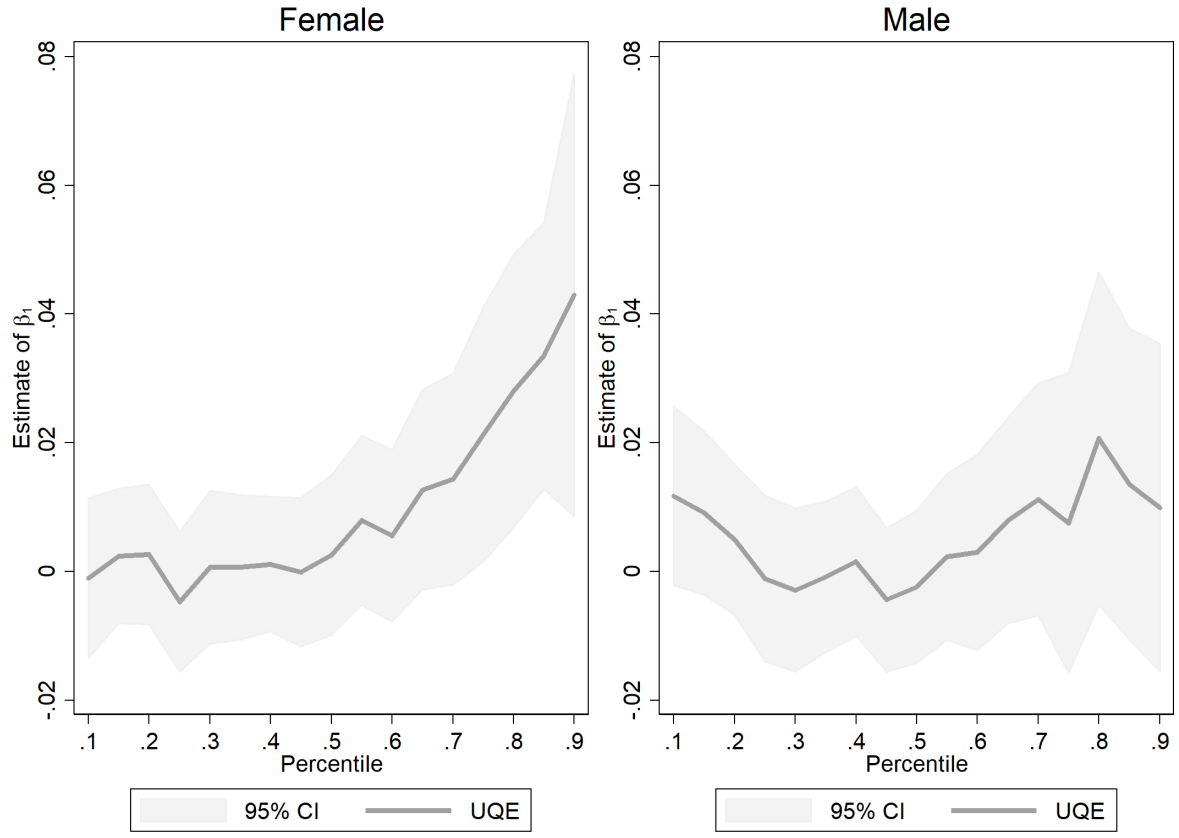
	(1)	(2)	(3)	(4)
Panel A: Males and Females				
Grademates BMI PGS (Normalized)	0.014*	0.014*	0.014*	0.012
	(0.007)	(0.007)	(0.007)	(0.008)
BMI PGS (Normalized)	0.065***	0.055***	0.056***	0.054***
	(0.007)	(0.007)	(0.007)	(0.011)
Observations	4,450	4,450	4,450	4,450
R-squared	0.088	0.130	0.132	0.136
Panel B: Females				
Grademates BMI PGS (Normalized)	0.028***	0.028***	0.028***	0.028***
	(0.009)	(0.009)	(0.009)	(0.010)
BMI PGS (Normalized)	0.071***	0.062***	0.061***	0.060***
	(0.010)	(0.010)	(0.010)	(0.018)
Observations	2,343	2,343	2,343	2,343
R-squared	0.149	0.176	0.177	0.180
Panel C: Males				
Grademates BMI PGS (Normalized)	0.004	0.001	0.000	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)
BMI PGS (Normalized)	0.059***	0.048***	0.049***	0.045***
	(0.010)	(0.009)	(0.009)	(0.017)
Observations	2,107	2,107	2,107	2,107
R-squared	0.110	0.171	0.176	0.184
Individual controls	Yes	Yes	Yes	Yes
Family and parental controls		Yes	Yes	Yes
Principal Components of SNP matrix			Yes	Yes
School-grade level controls				Yes

Note: The dependent variable is the obesity indicator measured at Wave II of Add Health. All regressions include school and grade fixed effects. Individual controls include age, age squared and gender and race indicators. Family and parental controls include family income, parental education, socio-economic status, parental obesity, the number of siblings and an indicator of whether both parents live in the household. Column (3) includes the first 10 principal components of the full genetic data. Column (4) includes average age, the share of females and blacks, grade size, average parental education - all computed at school-grade level - and average schoolmates BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

We also explore whether the effects of grade-mates' BMI PGS on individual BMI varies across its distribution by estimating unconditional quantile regressions. Table B.3 in Appendix

B and Figure 2 show that social-genetic effects are not statistically significant for males all along their $log(BMI)$ distribution. They are small and statistically insignificant for females in the lower to median part of the distribution and become increasingly larger and statistically significant from the 7th decile onwards.

Figure 2: The short-run (Wave II) effect of grade-mates' BMI PGS along the distribution of log BMI



Note: The figure shows the estimated unconditional quantile effects of grade-mates' BMI PGS on individual BMI and their associated confidence intervals. Dependent variable: $log(BMI)$ measured in Wave II. See the note to Table 4 for the full list of controls included in the regression. Number of observations: 6030.

Our baseline specification assumes that social-genetic effects are linear. This may be restrictive. We therefore consider also a specification that replaces grade-mates' average BMI PGS with the shares of grade-mates with a high (above the 75th percentile) and low BMI PGS (below the 25th percentile of the distribution in the grade and school). Panel A of Table 5 reports the results for obesity measured in Wave II. We find that the positive effect of grade-mates' BMI PGS on female obesity shown in Table 4 is driven by the share of peers with high BMI PGS, and confirm our baseline effects for males (no statistically significant effect). We estimate that a ten percentage point increase in the share of grade-mates with a high BMI PGS increases the probability that females are obese by 2.63 percentage points (a 20.9 percent increase with respect to the average probability of obesity). On the other hand, a similar in-

Table 5: The effect of the distribution of grade-mates' BMI PGS on obesity in Waves II (short-run) and IV (long-run)

	Males and Females	Females	Males
Panel A: Obesity Wave II			
Share of grademates with BMI PGS<25 pct.	-0.016	0.003	-0.016
	(0.099)	(0.116)	(0.155)
Share of grademates with BMI PGS>75 pct.	0.154	0.263**	0.056
	(0.104)	(0.110)	(0.160)
BMI PGS (Normalized)	0.054***	0.061***	0.045**
	(0.011)	(0.018)	(0.017)
Observations	4,450	2,343	2,107
R-squared	0.137	0.180	0.184
Panel B: Obesity Wave IV			
Share of grademates with BMI PGS<25 pct.	0.079	0.026	0.109
	(0.092)	(0.115)	(0.142)
Share of grademates with BMI PGS>75 pct.	-0.004	-0.035	0.046
	(0.095)	(0.132)	(0.136)
BMI PGS (Normalized)	0.098***	0.080***	0.118***
	(0.015)	(0.025)	(0.020)
Observations	5,957	3,139	2,818
R-squared	0.108	0.154	0.147

Note: The dependent variables are measured in Waves II and IV of Add Health. All regressions include school and grade fixed effects, age, age squared, gender and race indicators, family income, parental education, socio-economic status, parental obesity, the number of siblings, an indicator of whether both parents live in the household and the first 10 principal components of the full genetic data. The following grade-mates' controls are included: average age, the share of females and blacks, grade size, average parental education — all computed at school-grade level — and average schoolmates' BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

crease in the share of peers with low genetic predisposition to obesity has a negligible and statistically insignificant effect.

4.2 Long-run effects

A key question is whether the short-run effects of being exposed to high school grade-mates with higher genetic predisposition to BMI that we have uncovered for females persist into adulthood. We address this question by looking at obesity in Wave IV – when individuals are on average 12.8 years older than in Wave I.

Our results are summarized in Table 6. We find that the estimated long run social-genetic effect on obesity —captured by the estimated coefficient of grade-mates' BMI PGS– is nega-

Table 6: The long-run (Wave IV) effect of grade-mates' and own BMI PGS on obesity

	Males and Females	Females	Males
Grademates BMI PGS (Normalized)	-0.014*	-0.008	-0.017
	(0.008)	(0.013)	(0.012)
BMI PGS (Normalized)	0.098***	0.080***	0.118***
	(0.015)	(0.025)	(0.020)
Observations	5,957	3,139	2,818
R-squared	0.109	0.154	0.148

Note: The dependent variable is the obesity indicator measured in Wave IV of Add Health. All regressions include school and grade fixed effects, age, age squared, gender and race indicators, family income, parental education, socio-economic status, parental obesity, the number of siblings, an indicator of whether both parents live in the household and the first 10 principal components of the full genetic data. The following grade-mates controls are included: average age, the share of females and blacks, grade size, average parental education —all computed at school-grade level— and average schoolmates BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights are used. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

tive, close to zero, and statistically insignificant for both males and females, indicating that the positive influence of high school grade-mates on adolescent obesity dissipates over time. While the social effect dies out with time, the effect of own BMI PGS increases in importance as time goes by.

As shown by Table B.4 in Appendix B, long-run social-genetic effects are small and statistically insignificant not only at the mean but also along the log (BMI) distribution. In addition, Panel B of Table 5 indicates that in the long-run obesity is not affected significantly by the share of grade-mates who have either a high or a low BMI PGS.

4.3 The Effect of Grade-mates' Obesity on Individual Obesity.

Although we do not directly estimate the effect of grade-mates' obesity on the individual probability of being obese, we can derive this key effect from our estimates of individual obesity on own and grademates' BMI PGS. Under the assumption that the number of students in each grade within a school is sufficiently large, we show in Appendix A that this effect is equal to $\frac{\beta_1}{\beta_1+\beta_2}$. Table 7 reports the short-run and long-run values of $\frac{\beta_1}{\beta_1+\beta_2}$ using the estimates presented in Tables 4 and 6. We find that a ten percentage point increase in the share of obese grade-mates significantly increases —in the short-run— the probability that females are obese by 3.2 percentage points, but has no effects on males. In the long-run, however, this effect disappears.

4.4 Mechanisms

We have found evidence of a statistically significant short-run relationship between high school grade-mates' average BMI PGS and individual obesity, but only for females. In this

Table 7: The effect of grade-mates' obesity on individuals' obesity in Waves II (short-run) and IV (long-run)

	Males and Females	Females	Males
Panel A: Obesity W2			
Share of obese grade-mates	0.185*	0.321***	-0.101
	(0.103)	(0.102)	(0.280)
Observations	4,450	2,343	2,107
Panel B: Obesity W4			
Share of obese grade-mates	-0.165	-0.113	-0.172
	(0.118)	(0.204)	(0.139)
Observations	5,957	3,139	2,818

Note: The estimated effect of grade-mates obesity is computed as $\frac{\beta_1}{\beta_1+\beta_2}$, using the estimates of β_1 and β_2 from (2) reported in Tables 4 and 6, respectively. Derivation details are provided in Appendix A. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$

section, we discuss distorted perception of body size, nutrition and exercising choices as candidate mechanisms for the uncovered effects.

Although individuals do not observe genes directly, they observe peers' characteristics —including nutrition patterns, willingness to exercise and group obesity— that are affected by peers' BMI PGS. Social interactions and interpersonal comparisons may alter individuals' perceptions of body size and behaviours. This might happen because exposure to grade-mates with higher (lower) BMI may put individuals under less (more) pressure to conform to an ideal "thin body" norm or standard. In other words, individual identity or sense of belonging to a social category affects behaviour because different social categories are associated with different ideal physical attributes and behavioural prescriptions. Violating these prescriptions and deviating from these ideal attributes may reduce utility.[35] Alternatively, individuals can alter their behaviour by observing and imitating peers' behaviours (examples include eating habits and exercising) even in the absence of social pressure (Bandura and Walters, 1977; Chartrand and Bargh, 1999).

4.4.1 Distorted perception of body size

Interpersonal comparisons may result in distorted body-size perceptions during adolescence, which may in turn explain why females who are exposed to high school grade-mates with lower genetic predisposition to high BMI are less likely to be obese. To test this hypothesis, we investigate the relationship between peers' average BMI PGS and the discrepancy between perceived and actual body size. Our measure of perceived body size is based on answers to the following question to Wave II respondents: "How do you think of yourself in terms of

[35]See the theoretical contribution by Akerlof and Kranton (2000) for an introduction of identity into economic analysis, and Blanchflower et al. (2009) for an exploration of the importance of BMI comparisons for the spread of obesity.

weight?

1. very underweight
2. slightly underweight
3. about the right weight
4. slightly overweight
5. very overweight"

We group individual weight perceptions into four categories: (i) underweight, if respondents answer (1) or (2); (ii) normal, if they answer (3); (iii) overweight, if they answer (4); and (iv) obese, if they answer (5). We compare perceived with actual BMI and construct binary indicators identifying individuals who overestimate (underestimate) their BMI.

Columns 1 and 2 of Table 8 show the effect of grade-mates' average BMI PGS —computed in Wave I— and own BMI PGS on the probability that individuals underestimate or overestimate their weight in Wave II. We find that the individuals in our sample are more likely to underestimate their weight and less likely to overestimate it when their grade-mates' average BMI PGS is higher (although these effects are statistically significant only at the 10% level of confidence when we consider males and females separately). These results are consistent with Arduini et al. (2019), who find that the probability of reporting an upward (downward) distorted body-size perception is negatively (positively) correlated with peers' BMI.

4.4.2 Nutrition

Social-genetic effects may influence individual obesity via nutrition habits. If those with high BMI PGS are more prone to consuming sugar and calorie-dense foods, having high school grade-mates with high average BMI PGS may induce individuals to make similar choices, either because of imitation or because of distorted perceptions about their own BMI.[36]

We explore the role played by nutritional choices as a mechanism accounting for the effect of grade-mates' BMI PGS on individual obesity by considering the following Wave II questions: "Think about everything you had to eat and drink yesterday. This includes snacks as well as your regular meals.

1. Did you eat doughnuts, sweet rolls, muffins, or pastries?
2. Did you eat hot dogs or frankfurters?
3. Pizza?
4. French fries?
5. Potato chips, corn chips, tortilla chips, pretzels, or popcorn?

[36] In line with this, Arduini et al. (2019) find that peers' BMI affect the onset of eating disorders for females and that distorted body-size perception is an important mediator of this effect.

Table 8: Potential mechanisms accounting for the short-run (Wave II) effect of grade-mates and own BMI PGS on adolescents' obesity

Dependent Variable:	Underestimate weight	Overestimate weight	Unhealthy nutrition summary index	Exercising summary index	Passive activities summary index
Panel A: Males and Females					
Grademates BMI PGS (Normalized)	0.030***	-0.021**	0.032	-0.023	-0.013
	(0.010)	(0.009)	(0.021)	(0.025)	(0.024)
BMI PGS (Normalized)	-0.006	-0.011	-0.047	-0.049	0.034
	(0.019)	(0.017)	(0.045)	(0.039)	(0.036)
Observations	4,448	4,448	4,449	4,448	4,428
R-squared	0.095	0.092	0.116	0.151	0.138
Panel B: Females					
Grademates BMI PGS (Normalized)	0.024*	-0.025*	0.065**	0.014	0.022
	(0.012)	(0.013)	(0.026)	(0.037)	(0.029)
BMI PGS (Normalized)	0.021	-0.026	-0.032	-0.124**	0.007
	(0.021)	(0.026)	(0.060)	(0.050)	(0.049)
Observations	2,342	2,342	2,342	2,342	2,333
R-squared	0.110	0.100	0.113	0.200	0.129
Panel C: Males					
Grademates BMI PGS (Normalized)	0.033*	-0.017	0.002	-0.071**	-0.032
	(0.017)	(0.014)	(0.032)	(0.030)	(0.039)
BMI PGS (Normalized)	-0.035	-0.007	-0.057	0.007	0.071
	(0.031)	(0.013)	(0.084)	(0.073)	(0.061)
Observations	2,106	2,106	2,107	2,106	2,095
R-squared	0.105	0.114	0.170	0.155	0.174

Note: The dependent variables are measured in Wave II of Add Health. All regressions include school and grade fixed effects, age, age squared, gender and race indicators, family income, parental education, socio-economic status, parental obesity, the number of siblings, an indicator of whether both parents live in the household and the first 10 principal components of the full genetic data. The following grade-mates' controls are included: average age, the share of females and blacks, grade size, average parental education —all computed at school-grade level— and average schoolmates BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$

6. Cookies, brownies, cake, or pie?

7. Ice cream?

8. Yesterday, did you eat chocolate bars or candy?

9. Did you eat chicken or turkey? Was it fried?

10. Bacon, sausage, or chorizo?" [37]

Since the likelihood of rejecting a null hypothesis when it is true increases when multiple hypotheses are tested (Anderson, 2008), we create an index that combines these nutritional choices as follows (see Kling et al. 2007):[38]

$$Y^* = \frac{\sum_k Y_k^*}{K} \text{ , where } Y_k^* = \frac{Y_k - \mu_k}{\sigma_k}$$

, where Y_k^* is the kth component of the index, μ_k denotes its mean and σ_k its standard deviation. Since all the components are associated with higher propensity to consume sugar and calorie-dense food, higher values of the summary index Y^* are associated with higher propensity to consume unhealthy foods.

Column 3 of Table 8 shows the estimated effect of peers' average BMI PGS and own BMI PGS on this summary index. The results suggest that females are more likely to eat unhealthy foods when they have peers with higher average BMI PGS. For males, the relationship between nutrition and peers' BMI PGS is also positive but not statistically significant, in line with the fact that their obesity is not associated with peers' BMI PGS.

4.4.3 Physical Activity

Exercise is an additional potential mechanism. We construct a physical activity summary index using the following Wave II questions:

1. "During the past week, how many times did you go roller-blading, roller-skating, skate-boarding, or bicycling?

2. During the past week, how many times did you play an active sport, such as baseball, softball, basketball, soccer, swimming, or football?

3. During the past week, how many times did you do exercise, such as jogging, walking, karate, jumping rope, gymnastics or dancing?"

[37] We code the answers to question 9 "Did you eat chicken or turkey?" as positive only if the answer to the question "Was it fried?" was positive.

[38] Summary index tests have several advantages over testing individual outcomes and subsequently adjusting p-values to account for multiple inference. In our application, relying on a summary index allows us to test whether there is a "general" social genetic effect of interest on a family of outcomes (such as unhealthy foods consumption). Additionally, performing a summary index test is potentially more powerful than performing several individual-level tests because an index can reduce the amount of random error in each nutrition related outcome by combining data from multiple outcomes into a single index. See Anderson (2008) for further details.

Column 4 of Table 4.4 reports the estimated effect of grade-mates' BMI PGS on this index. We find no effect for the full sample and form females, and a negative (and statistically significant) effect for males. Although male teenagers exercise less when their grade-mates have higher average BMI PGS, this mechanism is not sufficient to alter their propensity to teenage obesity.

We also test whether grade-mates' BMI PGS is associated with an increase of sedentary activities (e.g., watching television), that we proxy with passive activities index, that summarizes the answers to the following Wave II questions:

1. "How many hours a week do you watch television?
2. How many hours a week do you watch videos?
3. How many hours a week do you play video or computer games?"

As shown in column 5 of Table 8, there is no evidence that this is the case.

5 Robustness checks

5.1 Attrition and the availability of valid genetic data

Due to the longitudinal nature of Add Health and to the fact that genetic data were collected at Wave IV, we lose about 20% of the sample because of attrition between Waves I and IV. Additionally, as described in Section 2.1, and despite the fact that Add Health obtained saliva samples for most of Wave IV respondents, non-eligibility for genotyping and the application of quality control procedures further reduced our working sample.

The validity of our results may be questioned if attrition and/or the absence of valid genetic information are systematically correlated with grade-mates' average BMI PGS. To evaluate this threat, we regress binary variables identifying missing observations due to: (i) attrition between waves I and IV; (ii) absence of valid genetic data (conditional on the individual being interviewed at Wave IV); (iii) either of the two reasons, on grade-mates' average BMI PGS, school and grade fixed effects. Reassuringly, we always find that the estimated grade-mates' effects are never statistically significant, suggesting that missing data are uncorrelated with our key explanatory variable.[39]

5.2 Results based on Probit Estimates

We have used a linear probability model to estimate the effects of peers' BMI PGS on individual obesity. In this sub-section, we verify whether our estimates vary when using a Probit specification. Results in Table B.5 in Appendix B show that the estimated average partial effect for females, which is statistically significant at the 1 percent, is slightly larger with the Probit model than with the linear probability model. The effect for males is very similar across models —it is not statistically significant and very close to zero.

[39] These results are available upon request from the authors.

6 Conclusion

We have used data from three waves of Add Health to study the short and long-run effects of high school grade-mates' genetic predisposition to high BMI —measured by average BMI polygenic scores— on adolescent and adult obesity. We have found that, in the short-run, a one standard deviation increase in these scores raises the probability of obesity for females by 2.8 percentage points, about half the size of the effect induced by a one standard deviation increase in the own polygenic score. Therefore, while adolescent obesity is partly in the genes, it is also in the genes of individuals one interacts with, especially for teenage girls, who are more vulnerable to peers than teenage boys.

In the long-run, however, our evidence indicates that the social-genetic effect induced by interactions with high school grade-mates fades away, suggesting that the relevant peer group for adults may shift from grade-mates to other peers, including co-workers.

Our results indicate that, for females, omitting social-genetic effects under-estimates the impact of genes on average obesity (in the school and grade). They also suggest that the negative consequences of a relatively poor genetic draw (for instance a BMI PGS that is 0.1 standard deviations above the average) on obesity can be offset during adolescence by interacting with a group of grade-mates with average BMI PGS about 0.2 standard deviations below the average. Even assuming that policies could be designed that affect the selection of grade-mates so as to compensate poor genetic draws, these policies would have an effect on adolescent obesity but not on adult obesity.

References

Akerlof, G. A. and Kranton, R. E. (2000): "Economics and identity," *The Quarterly Journal of Economics*, 115, 715–753.

Anderson, M. L. (2008): "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 103, 1481–1495.

Arduini, T., Iorio, D. and Patacchini, E. (2019): "Weight, reference points, and the onset of eating disorders," *Journal of health economics*, 65, 170–188.

Argys, L. M. and Rees, D. I. (2008): "Searching for peer group effects: A test of the contagion hypothesis," *The Review of Economics and Statistics*, 90, 442–458.

Asirvatham, J., Thomsen, M. R., Nayga Jr, R. M. et al. (2018): "Do peers affect childhood obesity outcomes? Peer-effect analysis in public schools," *Canadian Journal of Economics/Revue canadienne d'économique*, 51, 216–235.

Bandura, A. and Walters, R. H. (1977): *Social learning theory*, vol. 1, Prentice-hall Englewood Cliffs, NJ.

Barcellos, S. H., Carvalho, L. S. and Turley, P. (2018): "Education can reduce health disparities related to genetic risk of obesity: Evidence from a British reform," *BioRxiv*, 260463.

Barth, D., Papageorge, N. W. and Thom, K. (2019): "Genetic Endowments and Wealth Inequality," *Journal of Political Economy (forthcoming)*.

Beauchamp, J. P., Cesarini, D., Johannesson, M. et al. (2011): "Molecular Genetics and Economics," *Journal of Economic Perspectives*, 25, 57–82.

Belsky, D. W., Domingue, B. W., Wedow, R. et al. (2018): "Genetic analysis of social-class mobility in five longitudinal studies," *Proceedings of the National Academy of Sciences*, 115, E7275–E7284.

Benjamin, D. J., Cesarini, D., Chabris, C. F. et al. (2012): "The promises and pitfalls of genoeconomics," *Annu. Rev. Econ.*, 4, 627–662.

Bertoni, M., Brunello, G. and Cappellari, L. (2017): "Parents, Siblings and Schoolmates: The Effects of Family-School Interactions on Educational Achievement and Long-Term Labor Market Outcomes," .

Bierut, L., Biroli, P., Galama, T. et al. (2018): "Childhood socioeconomic status moderates genetic predisposition for peak smoking," *bioRxiv*, 336834.

Bifulco, R., Fletcher, J. M. and Ross, S. L. (2011): "The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health," *American Economic Journal: Economic Policy*, 3, 25–53.

Biroli, P. (2015): "Genetic and economic interaction in the formation of human capital: the case of obesity," Tech. rep., Mimeo. University of Zurich.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2013): "Under pressure? The effect of peers on outcomes of young adults," *Journal of Labor Economics*, 31, 119–153.

Blanchflower, D. G., Van Landeghem, B. and Oswald, A. J. (2009): "Imitative obesity and relative utility," *Journal of the European Economic Association*, 7, 528–538.

Böckerman, P., Cawley, J., Viinikainen, J. et al. (2019): "The effect of weight on labor market outcomes: An application of genetic instrumental variables," *Health economics*, 28, 65–77.

Bolyard, A. and Savelyev, P. A. (2019): "Understanding the Education Polygenic Score and Its Interactions with SES in Determining Health in Young Adulthood," *Available at SSRN 3397735*.

Breinholt, A. and Conley, D. (2019): "Child-driven parenting: differential early childhood investmnt by offsrping genotype," Mimeo, Available at http://paa2019.populationassociation.org/uploads/191088.

Brunello, G. and d'Hombres, B. (2007): "Does body weight affect wages?: Evidence from Europe," *Economics & Human Biology*, 5, 1–19.

Brunello, G., Michaud, P.-C. and Sanz-de Galdeano, A. (2008): "The rise in obesity across the Atlantic: An economic perspective," .

Carrell, S. E., Hoekstra, M. and West, J. E. (2011): "Is poor fitness contagious?: Evidence from randomly assigned friends," *Journal of public Economics*, 95, 657–663.

Cawley, J. (2004a): "An economic framework for understanding physical activity and eating behaviors," *American journal of preventive medicine*, 27, 117–125.

——— (2004b): "The impact of obesity on wages," *Journal of Human resources*, 39, 451–474.

——— (2011): *The Oxford handbook of the social science of obesity*, Oxford University Press.

——— (2015): "An economy of scales: A selective review of obesity's economic causes, consequences, and solutions," *Journal of health economics*, 43, 244–268.

Cawley, J., Han, E., Kim, J. J. et al. (2017): "Testing for peer effects using genetic data," Tech. rep., National Bureau of Economic Research.

Cawley, J., Han, E. and Norton, E. C. (2011): "The validity of genes related to neurotransmitters as instrumental variables," *Health economics*, 20, 884–888.

Chartrand, T. L. and Bargh, J. A. (1999): "The chameleon effect: the perception–behavior link and social interaction." *Journal of personality and social psychology*, 76, 893.

Cohen-Cole, E. and Fletcher, J. M. (2008): "Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic," *Journal of health economics*, 27, 1382–1387.

Conley, D. and Fletcher, J. (2018a): *The Genome Factor: What the social genomics revolution reveals about ourselves, our history, and the future*, Princeton University Press.

——— (2018b): *The Genome Factor: What the social genomics revolution reveals about ourselves, our history, and the future*, Princeton University Press.

Cools, A., Fernández, R. and Patacchini, E. (2019): "Girls, Boys, and High Achievers," Tech. rep., National Bureau of Economic Research.

Cornelissen, T., Dustmann, C. and Schönberg, U. (2017): "Peer effects in the workplace," *American Economic Review*, 107, 425–56.

Crosnoe, R., Frank, K. and Mueller, A. S. (2008): "Gender, body size and social relations in American high schools," *Social Forces*, 86, 1189–1216.

Cutler, D. M., Glaeser, E. L. and Shapiro, J. M. (2003): "Why have Americans become more obese?" *Journal of Economic perspectives*, 17, 93–118.

De La Haye, K., Robins, G., Mohr, P. et al. (2011): "Homophily and contagion as explanations for weight similarities among adolescent friends," *Journal of Adolescent Health*, 49, 421–427.

Dijkstra, P., Kuyper, H., Van der Werf, G. et al. (2008): "Social comparison in the classroom: A review," *Review of educational research*, 78, 828–879.

Domingue, B. W. and Belsky, D. W. (2017): "The social genome: Current findings and implications for the study of human genetics," *PLoS genetics*, 13, e1006615.

Domingue, B. W., Belsky, D. W., Conley, D. et al. (2015): "Polygenic influence on educational attainment: New evidence from the National Longitudinal Study of Adolescent to Adult Health," *AERA open*, 1.

Domingue, B. W., Belsky, D. W., Fletcher, J. M. et al. (2018): "The social genome of friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health," *Proceedings of the National Academy of Sciences*, 115, 702–707.

Fletcher, J. M. (2011): "Peer effects and obesity," *The Oxford handbook of the social science of obesity*, 303–312.

Fletcher, J. M., Ross, S. L. and Zhang, Y. (2013): "The determinants and consequences of friendship composition," Tech. rep., National Bureau of Economic Research.

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

Hoxby, C. (2000): "Peer effects in the classroom: Learning from gender and race variation," Tech. rep., National Bureau of Economic Research.

Hussey, J. M., Nguyen, Q. C., Whitsel, E. A. et al. (2015): "The reliability of in-home measures of height and weight in large cohort studies: Evidence from add health," *Demographic research*, 32, 1081.

Jencks, C. (1980): "Heredity, environment, and public policy reconsidered," *American Sociological Review*, 723–736.

Jou, C. (2014): "The biology and genetics of obesity—a century of inquiries," *New England Journal of Medicine*, 370, 1874–1877.

Kling, J. R., Liebman, J. B. and Katz, L. F. (2007): "Experimental analysis of neighborhood effects," *Econometrica*, 75, 83–119.

Kong, A., Thorleifsson, G., Frigge, M. L. et al. (2018): "The nature of nurture: Effects of parental genotypes," *Science*, 359, 424–428.

Kuczmarski, R. J. (2002): "2000 CDC growth charts for the United States; methods and development," .

Lavy, V. and Schlosser, A. (2011): "Mechanisms and impacts of gender peer effects at school," *American Economic Journal: Applied Economics*, 3, 1–33.

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

Locke, A. E., Kahali, B., Berndt, S. I. et al. (2015): "Genetic studies of body mass index yield new insights for obesity biology," *Nature*, 518, 197.

Loh, C.-P. A. and Li, Q. (2013): "Peer effects in adolescent bodyweight: Evidence from rural China," *Social Science & Medicine*, 86, 35–44.

Manski, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 60, 531–542.

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

Papageorge, N. W. and Thom, K. (2019): "Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study," *Journal of the European Economic Association (forthcoming)*.

Price, A. L., Patterson, N. J., Plenge, R. M. et al. (2006): "Principal components analysis corrects for stratification in genome-wide association studies," *Nature genetics*, 38, 904.

Reilly, J. J. and Kelly, J. (2011): "Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review," *International journal of obesity*, 35, 891.

Rodríguez-Planas, N., Sanz-de Galdeano, A. and Terskaya, A. (2018): "Independent Thinking and Hard Working, or Caring and Well Behaved? Short-and Long-Term Impacts of Gender Identity Norms," .

Ronda, V., Agerbo, E., Bleses, D. et al. (2019): "Family Disadvantage, Gender and the Returns to Genetic Human Capital," Working paper.

Sandholt, C., Hansen, T. and Pedersen, O. (2012): "Beyond the fourth wave of genome-wide obesity association studies," *Nutrition & diabetes*, 2, e37.

Sanz-de Galdeano, A. and Terskaya, A. (2019): "Sibling Differences in Educational Polygenic Scores: How Do Parents React?" .

Schaefer, D. R. and Simpkins, S. D. (2014): "Using social network analysis to clarify the role of obesity in selection of adolescent friends," *American journal of public health*, 104, 1223–1229.

Simmonds, M., Llewellyn, A., Owen, C. et al. (2016): "Predicting adult obesity from childhood obesity: a systematic review and meta-analysis," *Obesity reviews*, 17, 95–107.

Sotoudeh, R., Conley, D. and Harris, K. M. (2017): "The Influence of peer genotypes and behavior on smoking outcomes: Evidence from Add Health," Tech. rep., National Bureau of Economic Research.

Sotoudeh, R., Harris, K. M. and Conley, D. (2019): "Effects of the peer metagenomic environment on smoking behavior," *Proceedings of the National Academy of Sciences*, 201806901.

Speliotes, E. K., Willer, C. J., Berndt, S. I. et al. (2010): "Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index," *Nature genetics*, 42, 937.

Trogdon, J. G., Nonnemaker, J. and Pais, J. (2008): "Peer effects in adolescent overweight," *Journal of health economics*, 27, 1388–1399.

Wertz, J., Moffitt, T. E., Agnew-Blais, J. et al. (2019): "Using DNA from mothers and children to study parental investment in children's educational attainment," *Child Development*, forthcoming.

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

Yang, W., Kelly, T. and He, J. (2007): "Genetic epidemiology of obesity," *Epidemiologic reviews*, 29, 49–61.

A Appendix

In this Appendix we show that our estimates of the the effect of grade-mates' average PGS are informative on the relationship between individuals' and their peers' obesity.

Our baseline equation is (omitting controls for simplicity):

$$O_{isg} = \beta_0 + \beta_1 \overline{PGS}_{-i,sg} + \beta_2 PGS_{isg} + \epsilon_{igs} \tag{A.1}$$

where O stands for obesity, PGS_{isg} denotes individual i's own BMI PGS, and $\overline{PGS}_{-i,sg}$ is the average BMI PGS in the school and grade computed excluding individual i. Equation (A.1) holds for all grade-mates.

By averaging by school and grade we obtain

$$\overline{O}_{sg} = \beta_0 + (\beta_1 + \beta_2)\overline{PGS}_{sg} \tag{A.2}$$

where $\overline{O}_{sg}$ is the share of obese students in school s and g, and $\overline{PGS}_{sg}$ is the average BMI PGS in school s and grade g.

Therefore:

$$\overline{PGS}_{sg} = \frac{\overline{O}_{sg}}{\beta_1 + \beta_2} - \frac{\beta_0}{\beta_1 + \beta_2} \tag{A.3}$$

Note that:

$$\overline{PGS}_{-i,sg} = \frac{N_{sg}}{N_{sg} - 1}\overline{PGS}_{sg} - \frac{PGS_{isg}}{N_{sg} - 1} \tag{A.4}$$

where N_{sg} is the number of students in school s and grade g.

To write O_{isg} as a function of $\overline{PGS}_{sg}$ we substitute (A.4) into (A.1):

$$O_{isg} = \beta_0 + \left(\beta_2 - \frac{\beta_1}{N_{sg} - 1}\right) PGS_{isg} + \beta_1 \frac{N_{sg}}{N_{sg} - 1}\overline{PGS}_{sg} + \epsilon_{isg} \tag{A.5}$$

Finally, substituting (A.3) into (A.5) we obtain:

$$O_{isg} = \beta_0 - \frac{\beta_0 \beta_1}{\beta_1 + \beta_2}\frac{N_{sg}}{N_{sg} - 1} + \left(\beta_2 - \frac{\beta_1}{N_{sg} - 1}\right) PGS_{isg} + \frac{\beta_1}{\beta_1 + \beta_2}\frac{N_{sg}}{N_{sg} - 1}\overline{O}_{sg} + \epsilon_{isg} \tag{A.6}$$

Hence, when N_{sg} is large enough, (A.6) is approximately equivalent to:

$$O_{isg} = \beta_0 - \frac{\beta_0 \beta_1}{\beta_1 + \beta_2} + \beta_2 PGS_{isg} + \frac{\beta_1}{\beta_1 + \beta_2}\overline{O}_{sg} + \epsilon_{isg} \tag{A.7}$$

Therefore, the effect of peers' obesity on individual obesity is approximately equal to $\frac{\beta_1}{\beta_1 + \beta_2}$. Note that this results is only generalizable to linear-in-means reduced form models like (A.1), which include peers' average characteristics as an explanatory variable ($\overline{PGS}_{-i,sg}$ in our case).

B Data Appendix

Table B.1: Summary Statistics

	Males and Females		Females		Males	
	mean	st.dev.	mean	st.dev.	mean	st.dev.
Panel A: Individual Characteristics						
BMI PGS (Normalized)	0.000	1.000	0.025	0.987	-0.026	1.013
School grade 7	0.168	0.374	0.171	0.377	0.164	0.370
School grade 8	0.158	0.365	0.150	0.357	0.167	0.373
School grade 9	0.202	0.402	0.206	0.404	0.198	0.399
School grade 10	0.172	0.377	0.183	0.387	0.160	0.367
School grade 11	0.145	0.352	0.144	0.351	0.147	0.354
School grade 12	0.155	0.362	0.146	0.354	0.164	0.371
Age	15.946	1.730	15.842	1.701	16.053	1.754
Female	0.508	0.500	1.000	0.000	0.000	0.000
Black	0.150	0.357	0.162	0.369	0.137	0.344
White	0.753	0.431	0.750	0.433	0.757	0.429
Panel B: Family and Parental Characteristics						
Family income	0.429	0.395	0.430	0.396	0.428	0.394
Parents are college educated	0.193	0.394	0.191	0.393	0.194	0.395
SES index (Normalized)	-0.000	1.000	0.002	1.003	-0.002	0.997
At least one parent is obese	0.237	0.399	0.241	0.402	0.232	0.397
Number of siblings	1.512	1.347	1.505	1.406	1.519	1.283
Both parents live in household	0.713	0.453	0.706	0.456	0.720	0.449
Panel C: Grademates Characteristics						
Grademates BMI PGS (Normalized)	-0.000	1.000	0.002	0.999	-0.002	1.002
Average parental education of grademates	13.699	0.757	13.703	0.747	13.695	0.767
Share of black students	0.173	0.229	0.178	0.234	0.168	0.223
Average age of grademates	14.831	1.585	14.811	1.573	14.851	1.598
Share of female students	0.504	0.044	0.507	0.043	0.502	0.044
Grade size	224.140	128.432	221.939	125.331	226.412	131.537
Observations	6030		3176		2854	

Note: Means are computed using Wave I weights. All non-genetic variables are measured using Wave I of Add Health. Genetic variables are measured using Wave IV of Add Health.

Table B.2: The short-run (Wave II) effect of grade-mates' and own BMI PGS on adolescent obesity. Coefficient stability.

	(1)	(2)	(3)
Panel A: Males and Females			
Grademates BMI PGS (Normalized)	0.012		0.012
	(0.008)		(0.008)
BMI PGS (Normalized)		0.054***	0.054***
		(0.011)	(0.011)
Observations	4,450	4,450	4,450
R-squared	0.132	0.135	0.136
Panel B: Females			
Grademates BMI PGS (Normalized)	0.028***		0.028***
	(0.010)		(0.010)
BMI PGS (Normalized)		0.059***	0.060***
		(0.018)	(0.018)
Observations	2,343	2,343	2,343
R-squared	0.174	0.177	0.180
Panel C: Males			
Grademates BMI PGS (Normalized)	-0.004		-0.004
	(0.010)		(0.010)
BMI PGS (Normalized)		0.045***	0.045***
		(0.017)	(0.017)
Observations	2,107	2,107	2,107
R-squared	0.181	0.184	0.184

Note: The dependent variable is obesity measured in Wave II of Add Health. Columns (1) and (3) include average BMI PGS of grade-mates. Columns (2) -(3) include own BMI PGS. All regressions include school and grade fixed effects, age, age squared, gender and race indicators, family income, parental education, socio-economic status, parental obesity, the number of siblings, an indicator of whether both parents live in the household and the first 10 principal components of the full genetic data. The following grade-mates controls are included: average age, the share of females and blacks, grade size, average parental education —all computed at school-grade level— and average school-mates BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights.$***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$

Table B.3: The short-run (Wave II) effect of grade-mates' and own BMI PGS on adolescents' BMI along their BMI distribution. Unconditional Quantile Regression Estimates

	(1) 10th Centile	(2) 20th Centile	(3) 30th Centile	(4) 40th Centile	(5) 50th Centile	(6) 60th Centile	(7) 70th Centile	(8) 80th Centile	(9) 90th Centile
Panel A: Males and Females									
Grademates BMI PGS (Normalized)	0.003	0.003	-0.000	0.001	0.001	0.005	0.012*	0.020**	0.018*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.010)	(0.011)
BMI PGS (Normalized)	0.012	0.026***	0.028***	0.031***	0.031***	0.037***	0.041***	0.048***	0.059***
	(0.010)	(0.008)	(0.008)	(0.007)	(0.007)	(0.006)	(0.007)	(0.012)	(0.016)
Panel B: Females									
Grademates BMI PGS (Normalized)	-0.001	0.003	0.001	0.001	0.003	0.006	0.014*	0.028**	0.043**
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.007)	(0.008)	(0.011)	(0.017)
BMI PGS (Normalized)	0.020**	0.034***	0.036***	0.042***	0.036***	0.042***	0.039***	0.058***	0.100***
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.009)	(0.012)	(0.017)	(0.024)
Panel C: Males									
Grademates BMI PGS (Normalized)	0.012*	0.005	-0.003	0.002	-0.002	0.003	0.011	0.021	0.010
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)	(0.013)	(0.013)
BMI PGS (Normalized)	0.004	0.015	0.027**	0.020*	0.024**	0.034***	0.049***	0.048***	0.039**
	(0.014)	(0.010)	(0.012)	(0.011)	(0.010)	(0.010)	(0.012)	(0.017)	(0.016)

Note: The dependent variable is $log(BMI)$ measured in Wave II of Add Health. All regressions include school and grade fixed effects, age, age squared, gender and race indicators, family income, parental education, socio-economic status, parental obesity, the number of siblings, an indicator of whether both parents live in the household and the first 10 principal components of the full genetic data. The following grade-mates controls are included: average age, the share of females and blacks, grade size, average parental education —all computed at school-grade level— and average schoolmates BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights.$***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.4: The long-run (Wave IV) effect of grade-mates' and own BMI PGS on adolescents' BMI along their BMI distribution. Unconditional Quantile Regression Estimates

	(1) 10th Centile	(2) 20th Centile	(3) 30th Centile	(4) 40th Centile	(5) 50th Centile	(6) 60th Centile	(7) 70th Centile	(8) 80th Centile	(9) 90th Centile
Panel A: Males and Females									
Grademates BMI PGS (Normalized)	0.002	0.002	0.001	0.000	-0.001	-0.003	-0.003	0.007	0.001
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)	(0.010)
BMI PGS (Normalized)	0.029**	0.040***	0.055***	0.061***	0.066***	0.066***	0.066***	0.059***	0.023
	(0.014)	(0.013)	(0.012)	(0.011)	(0.011)	(0.009)	(0.010)	(0.010)	(0.016)
Panel B: Females									
Grademates BMI PGS (Normalized)	0.009	0.006	0.011	0.007	0.004	-0.001	0.012	0.022*	0.009
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.012)	(0.011)	(0.011)
BMI PGS (Normalized)	0.025	0.041**	0.053**	0.061***	0.076***	0.070***	0.057***	0.054***	0.005
	(0.018)	(0.019)	(0.021)	(0.019)	(0.016)	(0.018)	(0.018)	(0.017)	(0.022)
Panel C: Males									
Grademates BMI PGS (Normalized)	-0.002	-0.008	-0.009	-0.006	-0.006	-0.007	-0.004	-0.006	-0.002
	(0.010)	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)	(0.008)	(0.011)
BMI PGS (Normalized)	0.040***	0.053***	0.062***	0.063***	0.060***	0.057***	0.064***	0.069***	0.046**
	(0.015)	(0.014)	(0.012)	(0.012)	(0.011)	(0.012)	(0.010)	(0.012)	(0.021)

Note: The dependent variable is $log(BMI)$ measured in Wave IV of Add Health. All regressions include school and grade fixed effects, age, age squared, gender and race indicators, family income, parental education, socio-economic status, parental obesity, the number of siblings, an indicator of whether both parents live in the household and the first 10 principal components of the full genetic data. The following grade-mates controls are included: average age, the share of females and blacks, grade size, average parental education —all computed at school-grade level— and average schoolmates BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.5: The short-run (Wave II) effect of own and grade-mates' BMI PGS on adolescents' obesity. LPM vs. Probit

	LPM	Probit
Panel A: Males and Females		
Grademates BMI PGS (Normalized)	0.012	0.018**
	(0.008)	(0.009)
BMI PGS (Normalized)	0.054***	0.056***
	(0.011)	(0.009)
Panel B: Females		
Grademates BMI PGS (Normalized)	0.028***	0.038***
	(0.010)	(0.010)
BMI PGS (Normalized)	0.060***	0.059***
	(0.018)	(0.015)
Panel C: Males		
Grademates BMI PGS (Normalized)	-0.004	-0.001
	(0.010)	(0.014)
BMI PGS (Normalized)	0.045***	0.051***
	(0.017)	(0.017)

Note: Average marginal effects are reported. The dependent variables are measured at Wave II of Add Health. All regressions include school and grade fixed effects, age, age squared, gender and race indicators, family income, parental education, socio-economic status, parental obesity, the number of siblings, an indicator of whether both parents live in the household and the first 10 principal components of the full genetic data. The following grade-mates controls are included: average age, the share of females and blacks, grade size, average parental education —all computed at school-grade level— and average schoolmates BMI PGS computed at the school level. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave I weights. Number of observations: 4,450. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$