

# Testing for family influences on obesity: The role of genetic nurture

John Cawley<sup>1</sup>  | Euna Han<sup>2</sup>  | Jiyeon Kim<sup>3</sup> | Edward C. Norton<sup>4</sup> 

<sup>1</sup>Department of Policy Analysis and Management, Cornell University and NBER, Ithaca, New York

<sup>2</sup>College of Pharmacy, Yonsei Institute of Pharmaceutical Science, Yonsei University, Incheon, South Korea

<sup>3</sup>Department of Economics, Elon University, Elon, North Carolina

<sup>4</sup>Department of Health Management and Policy and Department of Economics, University of Michigan and NBER, Ann Arbor, Michigan

## Correspondence

Edward C. Norton, University of Michigan and NBER, Ann Arbor, MI.

Email: ecnorton@umich.edu  
John Cawley, Cornell University and NBER, 2312 MVR Hall, Ithaca, NY 14853.

Email: jhc38@cornell.edu

## Funding information

Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: P01-HD31921

## Abstract

A large literature has documented strong positive correlations among siblings in health, including body mass index (BMI) and obesity. This paper tests whether that is explained by a specific type of peer effect in obesity: genetic nurture. Specifically, we test whether an individual's weight is affected by the genes of their sibling, controlling for the individual's own genes. Using genetic data in Add Health, we find no credible evidence that an individual's BMI is affected by the polygenic risk score for BMI of their full sibling when controlling for the individual's own polygenic risk score for BMI. Thus, we find no evidence that the positive correlations in BMI between siblings are attributable to genetic nurture within families.

## KEYWORDS

family, genetic nurture, genetics, health, obesity, peer effects, sibling, body mass index (BMI)

## JEL CLASSIFICATION

I1; I12; I14; D1; J13

## 1 | INTRODUCTION

A substantial literature documents that there is a strong positive correlation between siblings in their health and health behaviors. Halliday and Mazumder (2017) estimate that the correlations among young siblings in health conditions range from 0.5 to 0.6. In U.S. data, the correlation between siblings in smoking is 0.355, drinking alcohol is 0.350, frequency of TV watching is 0.366, and frequency of exercise is 0.209 (Daw, Margolis, & Verdery, 2015), and the correlation in body mass index (BMI)<sup>1</sup> is 0.53 (Brown & Roberts, 2012). These have been interpreted as “high” levels of homophily among siblings in health behaviors (Daw, Margolis, & Verdery, 2015). In addition, numerous studies have examined potential spillovers across siblings related to substance use (Altonji, Cattani, & Ware, 2017), birth weight (Breining, Daysal, Simonsen, & Trandafir, 2015), iodine supplementation (Adhvaryu & Nyshadham, 2016), and illness symptoms (Ho, 2017).<sup>2</sup>

<sup>1</sup>BMI is calculated as weight in kilograms divided by height in meters squared.

<sup>2</sup>A large literature also examines how the health of a youth may spill over to affect sibling education; for example, Black et al. (2017), Parman (2015), Breining, Daysal, Simonsen, and Trandafir (2015), Almond, Edlund, and Palme (2009), Breining (2014), Fletcher and Wolfe (2008), Fletcher, Hair, and Wolfe (2012), Fletcher, Mailick, Song, and Wolfe (2013), and Fletcher, Vidal-Fernandez, and Wolfe (2018).

There are many possible explanations for sibling correlations in health and health behaviors. Biologically related siblings have correlated genetic endowments. Siblings may experience similar environments, what Manski (1993) calls correlated effects. Parents may reallocate resources among their children in response to health and health shocks, which could either increase or decrease equity in health (Griliches, 1979; Yi, Heckman, et al. 2015). Similarity among siblings could also arise because of economies of scale and parental learning by doing in the production of child health (Corman, Dave, & Reichman, 2017). Finally, there may be endogenous peer effects among siblings or from parents to children (Leibenstein, 1950; Manski, 1993, 2000).

One specific type of peer effect within families is *genetic nurture* (sometimes called indirect genetic effects), which is defined as one individual's genes affecting another individual's outcomes (Wolf, Brodie, Cheverud, Moore, & Wade, 1998). Kong et al. (2018) test for genetic nurture using data for an Icelandic sample, by testing whether parental alleles (i.e., specific forms of a gene) that were not transmitted to (i.e., not inherited by) the child are still associated with the child's outcomes. They find evidence of genetic nurture in education; specifically, they find that parental alleles associated with education that were not transmitted to the child have 29.9% the effect of the alleles that were transmitted to the child. In contrast, they find no evidence of genetic nurture between parents and children in BMI; nontransmitted parental alleles associated with BMI have no detectable effect on child BMI. Other evidence on this question comes from Richmond et al. (2017), who test for genetic nurture in BMI using data from the United Kingdom. Similar to the earlier study, they find no detectable effect of nontransmitted alleles associated with high maternal BMI on the BMI of the mother's children.

This paper contributes to the literature by testing for genetic nurture in BMI and obesity. Specifically, we test whether a youth's weight is affected by their sibling's polygenic score (PGS) for BMI, controlling for the person's own PGS for BMI. (PGSs are based on multiple genes; in our case, it is based on 97 DNA locations, which in turn are referred to as single-nucleotide polymorphisms or SNPs.) Whereas past studies (Kong et al., 2018; Richmond et al., 2017) have examined genetic nurture from parents to children, we examine genetic nurture among siblings. Earlier studies that fail to detect genetic nurture from parent to child are not informative about whether there may exist genetic nurture among siblings. For example, siblings may spend more time playing with each other than with a parent and thus may have greater influence on each other's physical activity. Likewise, youth may be more responsive to cues about diet and appearance from a sibling than a parent; that is, a sibling may be a more powerful role model than parents for diet, physical activity, and weight.

Another contribution of this paper is that it is the first to test for genetic nurture in BMI in the United States; past studies examined data from Iceland and the United Kingdom. Genetic nurture among families may vary based on culture and society; for example, school sports are more important in the U.S. than in Iceland or the United Kingdom, and school sports participation could be one way in which siblings influence each other's weight. Another important consideration is that the age-standardized mean BMI is higher in the United States (28.9) than in the United Kingdom (27.4) or Iceland (26.8; see NCD Risk Factor Collaboration, 2016; Table A4), and the United States has been described as having a particularly "obesogenic" environment (Brownell & Horgen, 2004; Nestle, 2002), and thus, the extent of genetic nurture in BMI could be different in the United States.

BMI and obesity are important outcomes to study (Cawley, 2015). Obesity is highly prevalent in the United States; currently, 39.6% of U.S. adults aged 20 years and older, and 18.5% of youths aged 2–19 years, are obese (Hales, Fryar, Carroll, Freedman, & Ogden, 2018).<sup>3</sup> Worldwide, between 1975 and 2014, mean BMI rose from 21.7 to 24.2 for men and from 22.1 to 24.4 for women (NCD Risk Factor Collaboration, 2016). Obesity is of public health interest because excess fat raises the risk of heart disease, type II diabetes, and cancer (Hu, 2008), and as a result, the United States spends \$315.8 billion per year treating obesity-related illness (Cawley, 2015), and obesity is one of the top preventable causes of death in the United States (Mokdad, Marks, Stroup, & Gerberding, 2004, 2005). If there exist family or peer effects in BMI, that may help explain the continued rise in obesity (Sassi, 2010) and may also imply that effective weight-reduction interventions could have beneficial spillover effects within families.

In this paper, we test for genetic nurture using data on the genotypes of full siblings in the National Longitudinal Survey of Adolescent to Adult Health (Add Health). Our analysis yields no detectable evidence of genetic nurture between siblings in BMI; this finding is robust across two measures of weight (BMI and an indicator for obesity) and across numerous subgroups and waves of data. The results are not simply statistically insignificant; the point estimates

<sup>3</sup>Among adults, obesity is defined as a BMI  $\geq 30$ . Among youths 19 and younger, obesity is defined based on weight-for-height thresholds that vary by sex and month of age (Institute of Medicine, 2012).

are also quite small. Thus, we find no evidence that the large, positive correlations in BMI between siblings are attributable to genetic nurture or peer effects between siblings.

## 2 | METHODS

This study exploits the natural experiment of Mendelian randomization: variation in the random draw of genes that each full sibling receives from the same two parents. We measure the genetic endowment relating to BMI using the PGS for BMI, which reflects the combined additive influence of SNPs associated with BMI across the entire genome. The specific PGS we use is based on a genome-wide association study (GWAS) of 332,206 individuals (Locke, Kahali, Berndt, et al., 2015).<sup>4</sup> The GWAS identified 97 SNPs robustly associated with BMI<sup>5</sup> that collectively explain 2.7% of the variance of BMI (Locke, Kahali, Berndt, et al., 2015).<sup>6</sup> The PGS is equal to the number of alleles associated with BMI, each weighted by its relative contribution to BMI.

The PGS has several advantages in this context. It has high explanatory power for a single variable and has good out-of-sample reliability. Also, alleles are determined prior to birth (at conception) and are largely immutable and thus are not affected by the actions or characteristics of one's siblings.<sup>7</sup> To clarify, gene expression can be affected by the environment (this is the subject of the field of epigenetics) but generally not one's alleles, and thus, one's PGS is exogenously determined. Variation in the PGS within full siblings is the result of their random draw of genes from the same two parents (Mendelian randomization).

For nomenclature, we follow the convention of previous studies (e.g., Cohen-Cole & Fletcher, 2008) and refer to the individual whose behavior may be affected as the *ego* and the individual who may be affecting them as the *alter*.

Our basic reduced-form model regresses the ego's weight ( $Y$ ) on the alter's PGS for BMI, the ego's own PGS for BMI, various other ego-level exogenous explanatory variables  $X$ , and indicators for survey wave:

$$Y_{it} = \beta_0 + \beta_1 \text{AlterPGS}_{it} + \beta_2 \text{EgoPGS}_{it} + \beta_3 X_{it} + \beta_4 PC_{it} + \beta_5 \text{Wave}_t + \epsilon_{it} \quad (1)$$

The dependent variable  $Y$  is either continuous BMI or an indicator variable for obesity. When BMI is the dependent variable, we estimate ordinary least squares models. When the indicator of obesity is the dependent variable, we estimate linear probability models.

The parameter of interest is  $\beta_1$ , the coefficient on the alter's PGS for BMI. Because the model controls for the ego's own PGS for BMI, we are testing whether the variance in genetic predisposition of the alter that is uncorrelated with that of the ego affects the ego's weight. In this way, we test for genetic nurture. To clarify, we do not examine whether specific SNPs inherited by the alter but not the ego affect the ego's weight but whether the aggregate measured genetic predisposition of the alter that is independent of that of the ego is correlated with the ego's weight.

The distribution of allele frequency is not random across race and ethnic groups (Conley, Laidley, Boardman, & Domingue, 2016; Guo, Liu, Wang, Shen, & Hu, 2015; Liu & Guo, 2015; Price et al., 2006). If one did not address this, it could cause omitted variable bias. A standard approach in the literature is to control for the principal components of the GWAS; by doing so, one controls for ethnic and racial variation across the genome (see, e.g., Price et al., 2006; Schmitz & Conley, 2018). We follow that approach; the regression model above controls for 10 principal components from the GWAS (noted as  $PC$  in Equation 1).

To avoid the selection bias that would result from studying a self-selected peer group, we study peer groups of full siblings (i.e., siblings with the same two biologic parents). Because each of the two full siblings received a draw of genes

<sup>4</sup>These individuals were of European ancestry, and thus, the PGS based on this sample may not be as predictive of BMI for individuals of other ancestries (see Braudt, 2018; Martin et al., 2017; Ware et al., 2017). We address this later by estimating models separately by race.

<sup>5</sup>The genetics literature has investigated whether the SNPs associated with high BMI are also associated with other characteristics; they find that they are robustly associated with the components of BMI (weight and height) and obesity-related illnesses such as diabetes, high blood pressure, and high cholesterol (which are part of the total effect of obesity) but not characteristics unrelated to obesity (Locke, Kahali, Berndt, et al., 2015; Speliotes, Willer, Berndt, et al., 2010). In addition, a study of genetic influences on BMI concluded that pleiotropy was unlikely for the alleles influencing BMI (Richmond et al., 2017). Even if some of the SNPs in our PGS that affect BMI also affect other characteristics, that is not a problem for our study—we are testing whether there exists genetic nurture among siblings in BMI. The mechanism through which such nurture may operate is not of importance; we are testing whether it exists at all through any mechanism.

<sup>6</sup>Although it is estimated that 40–70% of the variation across individuals in BMI is due to genetic factors, the vast majority of genetic variability in BMI has not yet been linked to individual SNPs; see, for example, Locke, Kahali, Berndt, et al. (2015).

<sup>7</sup>The alter's genes could be altered by retroviruses or replication errors, but these are assumed to be unrelated to ego weight.

from the same two parents, variation in the PGS of the alter, controlling for the PGS of the ego, reflects Mendelian randomization. We exclude twins from our sample because they may exhibit different patterns of genetic nurture than other types of siblings. Specifically, twins may actively seek to establish their own identities or more enthusiastically mimic each other. In addition, the model cannot be estimated for monozygotic (i.e., identical) twins because for them, ego and alter genes are perfectly collinear, so there is no independent variation in the alter PGS controlling for ego PGS.

One strength of the reduced-form approach is that it does not restrict the possible mechanisms for genetic nurture or peer effects. They may operate through various *family mechanisms* such as the observed sibling (the alter), another sibling that we do not observe, or the parents. And they may operate through various *behavioral mechanisms* such as calorie intake (diet), calorie expenditure (physical activity), or preferences about weight or appearance. Any and all such effects will be reflected in the reduced-form estimate and are of interest to us.

In the regression model (1), the vector  $X$  includes the following set of control variables: the respondent's age in years; sex; race (African-American, Hispanic, and other, with white as the reference); an indicator for whether the respondent lives with his or her parents; an indicator for whether the siblings appear to be cohabiting (see Section 3 for details); total household income (\$20K–\$40K; \$40K–\$50K; \$50K–\$75K; \$75K–\$99K; and >\$100K, with <\$20K as the reference); the respondent's highest grade completed (high school graduate, less than college, and bachelor's degree or more, with less than high school graduate as the reference); parental educational attainment (same educational categories); respondent's marital status (married with not married as the reference); and respondent's employment status (work full time and part-time, with unemployed or out-of-labor force as the reference). The identically independently distributed error term is  $\epsilon$ .

One might hypothesize that not all siblings may be equally influential. For example, Yakusheva, Kapinos, and Eisenberg (2014) found in their study of peer effects in weight among college roommates that higher status alters (defined as those who were in higher socioeconomic status or more sexually experienced) were more influential than lower status alters. Another example is that Carrell, Hoekstra, and West (2011) found, in their study of peer effects in fitness among Air Force cadets, that the least fit alters were most influential; the most fit did not exhibit a beneficial influence. Likewise, in our study, it may be that siblings do not affect each other equally. To investigate the possibility of heterogeneous effects, we estimate the regression model for different subsamples: (a) ego is older than the alter; (b) ego is younger than the alter; (c) siblings are the same sex; (d) siblings are of different sex; (e) ego is female; (f) ego is male; (g) ego is non-white; (h) ego is white; (i) ego has a PGS above the mean; and (j) ego has a PGS below the mean. When we test for heterogeneous treatment effects by subgroup, then we also control for the difference in age between the siblings and whether they are of the same sex.

We first estimate our models for three waves of data pooled. However, the association of certain genes with BMI can vary by age and time (e.g., see Rosenquist et al., 2015), so we also estimate our models for each wave separately. It is not possible to estimate an individual fixed effects model because the PGS for both egos and alters are time invariant.

Typically, each individual in the sample can appear in up to two observations per wave of the data: one in which they are the ego (i.e., their weight is the dependent variable) and one in which they are the alter (i.e., their PGS is the key independent variable). However, for the small number of households in which we observe three full siblings, individuals can appear in up to three observations per wave: one as the ego and two as the alter. To account for the correlation in error terms among siblings, we cluster the standard errors by family.

### 3 | DATA: NATIONAL LONGITUDINAL STUDY OF ADOLESCENT TO ADULT HEALTH

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a longitudinal study of a nationally representative sample of adolescents. Wave I of the study took place during the 1994–1995 school year when the subjects were age 12–19. Wave II surveyed the same adolescents a year later, in 1996. In Wave III, respondents were reinterviewed as they were entering young adulthood, aged 18–26, in 2001–2002. Wave IV followed in 2008–2009, when the subjects were aged 25–34 years.

The Add Health data are uniquely useful for this analysis because they contain a large number of siblings as well as genetic data. The design of Add Health includes an oversample of about 3,000 pairs of individuals with varying genetic relatedness, including twins, full siblings, half siblings, and adolescents with no biological relationship but who were raised in the same household. We identify sibling status among Add Health respondents using the sibling pair variable from Wave I.

In Wave IV (aged 25–34 years), trained field interviewers collected saliva samples from respondents who were members of sibling pairs,<sup>8</sup> which were then shipped to a lab for DNA extraction and genotyping (Smolen et al., 2013).<sup>9</sup> Roughly 96% consented to data collection and 78% to archiving of their data. The Add Health recently (in mid-2018) released the PGS for BMI. It is normalized within ancestry groups to have mean 0, standard deviation of 1.

We drop the sibling pairs if either of them has missing or invalid PGS for BMI. When data on control variables (other than PGS) are missing, the value is set to 0 and an indicator for missing values of that variable is set to 1 and included in the model.

Our main dependent variables are a linear measure of BMI and an indicator variable for obesity.<sup>10</sup> We calculate BMI using measurements (i.e., not self-reports) of weight and height, which avoids the complications associated with the reporting error that is common in weight (e.g., Cawley, Maclean et al., 2015). For our analyses, we pool together Waves II, III, and IV; measurements of weight and height were taken in each of those waves. We do not use data from Wave I because no measurements of weight and height were taken in that wave; only self-reported data were recorded. However, as an extension, we estimate models using the self-reported data in Wave I.

We control as well as possible, given the limitations of the data, for whether the siblings are cohabiting in the same household at the time of the interview. This information is not explicitly contained in the Add Health, so we impute it in the following manner. In Waves I, II, and IV, we assume that the ego and alter cohabit if the household roster indicates that the respondent is living with a sibling of a gender that matches the alter. In Wave III, we use the variable that asks about travel time to their sibling; we code the ego and alter as cohabiting if the respondent states that they live with their sibling. However, we cannot be sure that the respondent's sibling in this question is the sibling who is the alter in our model.

We estimate models with all waves pooled as well as separately by wave to examine whether effects vary by the age of the respondents. Genetic nurture effects could be stronger when people are younger and more likely to live together, or, because age affects gene expression, may be stronger later as the prevalence of obesity rises in later waves. The possibility that genetic nurture could vary with time is another reason for estimating the model separately by wave. The main sample for Waves II–IV has 2,546 observations on 914 unique individuals and 481 unique sibling pairs. When, as an extension, we add the wave I data to our sample, it contributes an additional 934 observations.

## 4 | EMPIRICAL RESULTS

### 4.1 | Summary statistics

Table 1 presents the summary statistics for the sample using all waves pooled. The mean BMI is 26.74, and 26.6% of the sample is obese. The mean PGS for BMI is 0.06 among the egos and 0.07 among the alters.

Although not shown in the table of summary statistics, the correlation in the PGS for BMI among the full siblings in our sample is 0.54. By all measures, the weight of siblings is highly correlated. The correlation in BMI is 0.43 for all waves pooled and is 0.42, 0.38, and 0.30 in Waves II, III, and IV, respectively. The sibling correlation in obesity status is 0.46 for all waves pooled and 0.45, 0.46, and 0.35 for Waves II, III, and IV, respectively. The correlation between siblings in waist circumference in Wave IV is 0.30. Thus, we confirm the consistent finding in the earlier literature that siblings have highly correlated BMI and obesity status. We next test whether this correlation is due in part to genetic nurture.

### 4.2 | Results of reduced-form regressions

Table 2 contains the results of our reduced-form regressions. Panel A is for all three waves (II, III, and IV) pooled, and panels B, C, and D are for each of those waves separately. Columns 1 and 2 report results for obesity, whereas columns 3

<sup>8</sup>Add Health genotyped and calculated BMI risk scores for roughly two thirds of its sibling pair subsample (McQueen et al., 2015). In Table A1, we conduct a balance test of the characteristics of Add Health respondents for whom the genetic risk score is available (1,283 individuals) and those for whom it is unavailable (4,388 individuals). Characteristics associated with the risk score being available (as opposed to missing) include race (valid risk score is more likely for African-Americans and less likely for other groups), a more probability that the sibling pairs live with parents, and a lower probability that parents are college graduates.

<sup>9</sup>These Add Health data have been used to study the genetics of BMI and obesity (e.g., Domingue et al., 2018; Haberstick et al., 2010).

<sup>10</sup>Respondents aged 20 years and older are classified as obese if their BMI equals or exceeds 30. Those aged 19 and younger are classified as obese using weight-for-height cutoffs that vary by sex and month of age (Institute of Medicine, 2012).

**TABLE 1** Summary statistics: Full siblings

Variable name	Mean	SD	Min	Max
Outcome of interest ( $N = 2,546$ )				
BMI	26.74	7.19	12.54	80.52
Obese	0.266	0.442	0	1
Ego characteristics ( $N = 2,546$ )				
PGS for BMI	0.0642	1.019	-2.855	3.508
PGS for waist circumference	0.0912	1.028	-3.032	3.462
Male	0.485	0.499	0	1
Age	22.64	5.305	12	34
African-American	0.155	0.362	0	1
Hispanic	0.085	0.278	0	1
Other race	0.092	0.289	0	1
White	0.667	0.472	0	1
Married	0.267	0.442	0	1
Less than HS	0.275	0.446	0	1
HS grad	0.331	0.471	0	1
Less than college	0.201	0.401	0	1
College grad	0.153	0.359	0	1
Working part-time	0.379	0.485	0	1
Working full time	0.509	0.500	0	1
Live with parents	0.476	0.500	0	1
Live with siblings	0.388	0.487	0	1
Alter characteristics ( $N = 2,546$ )				
PGS for BMI	0.067	1.018	-2.855	3.508
PGS for waist circumference	0.094	1.024	-3.032	3.462
Age difference (months)	26.59	11.39	9	69
Same-sex sibling	0.571	0.495	0	1
Household characteristics ( $N = 2,546$ )				
20K < income < 39K	0.199	0.399	0	1
40K < income < 49K	0.092	0.289	0	1
50K < income < 75K	0.183	0.386	0	1
75K < income < 99K	0.097	0.296	0	1
100K < income	0.092	0.288	0	1
Parents: less than HS	0.118	0.323	0	1
Parents: HS grad	0.307	0.461	0	1
Parents: less than college	0.254	0.435	0	1
Parents: college grad	0.208	0.406	0	1

Note. Data: Add Health Waves II, III, and IV, pooled. There are categories for “missing” for race, education, and income, which are included in the models but are not shown here.

Abbreviations: HS, high school; PGS, polygenic score.

and 4 report results for BMI. For each, we report results of models that exclude, and then include, the alter's PGS for BMI. In panel A, one can see that the ego's own PGS for BMI is strongly associated with a higher BMI for the ego. A one-standard deviation increase in the ego's PGS is associated with an increase in the probability of being obese by 6.84 percentage points (column 2) and with an increase in BMI of about 1.21 (column 4), which is roughly 7 to 9 lb for an adult of average height. In contrast, the PGS for BMI of the alter (their full sibling) has a much smaller point estimate and is not statistically significant. This is true for both obesity (column 2) and BMI (column 4). We can reject the null hypothesis that the effect sizes of ego PGS and alter's PGS are equal at  $p = .001$ . We cannot reject the null hypothesis that controlling for alter PGS has no impact on the magnitude of the coefficient on ego PGS ( $p = .1775$  for BMI and  $p = .4995$  for obesity).

We also estimate the model separately by wave of the Add Health because the effects of genetic nurture may vary with age or time. We find that the results by wave (panels B, C, and D for Waves II, III, and IV, respectively) exhibit

**TABLE 2** Regression results for Equation (1), full siblings only

<b>A. All waves pooled (Waves II, III, and IV)</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	26.59%		26.74	
Alter's PGS for BMI		−0.0085 [0.0127]		−0.2868 [0.2149]
Ego's PGS for BMI	0.0641*** [0.0128]	0.0684*** [0.0126]	1.0628*** [0.2121]	1.2091*** [0.2024]
Observations	2,546	2,546	2,546	2,546
R <sup>2</sup>	.123	.123	.213	.214
<b>B. Wave II</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	14.68%		23.30	
Alter's PGS for BMI		−0.0232 [0.0141]		−0.2657 [0.1826]
Ego's PGS for BMI	0.0671*** [0.0137]	0.0792*** [0.0149]	1.0701*** [0.1982]	1.2085*** [0.1890]
Observations	838	838	838	838
R <sup>2</sup>	.092	.095	.134	.135
<b>C. Wave III</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	24.34%		26.99	
Alter's PGS for BMI		−0.0102 [0.0172]		−0.2530 [0.2592]
Ego's PGS for BMI	0.0509*** [0.0179]	0.0560*** [0.0186]	0.9928*** [0.2638]	1.1180*** [0.2560]
Observations	756	756	756	756
R <sup>2</sup>	.101	.101	.101	.102
<b>D. Wave IV</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	38.87%		29.57	
Alter's PGS for BMI		0.0027 [0.0171]		−0.3453 [0.2801]
Ego's PGS for BMI	0.0676*** [0.0167]	0.0663*** [0.0170]	1.0649*** [0.2581]	1.2409*** [0.2649]
Observations	952	952	952	952
R <sup>2</sup>	.102	.102	.111	.113

*Note.* Robust standard errors are in brackets, clustered at the family level. First 10 principal components are included in the model. Demographic variables controlled for are age, sex, race, whether the respondent lives with parents, whether the siblings live together, total household income, the respondent's highest grade completed, marital status, employment status, and parental educational attainment. For the outcome of obese, the ordinary least squares (linear probability model) coefficient is reported.

Abbreviation: DV, dependent variable; PGS, polygenic score.

\*\*\* $p < .01$ . \*\* $p < .05$ . \* $p < .1$ .

the exact same pattern as the main results—ego PGS for BMI is consistently and strongly associated with a higher probability of obesity and a higher BMI of the ego, but the alter's PGS for BMI is essentially uncorrelated with ego obesity and BMI. The coefficients on alter PGS are not only not statistically significant, but they are also small in magnitude.

Guided by the literature on peer effects, which has at times found that some egos are more easily influenced than others and some alters are more influential than others (see, e.g., Yakusheva, Kapinos, & Eisenberg, 2014; Carrell, Hoekstra, & West, 2011), we next estimate models for different subgroups of siblings. One possible difference that could be important is relative age; for example, egos may be more influenced by alters who are older than them than by alters who are younger than them. In Table 3, panel A, we present results separately based on whether the ego is older than the alter or younger than the alter. In both cases, alter PGS for BMI is essentially uncorrelated with ego BMI and obesity, and the point estimates are not significantly different for alters who are younger versus those who are older. In other words, there is no evidence of genetic nurture no matter what the relative age of the ego to the alter.

Another potentially important factor is whether the ego and alter are of the same or different sex; for example, people may be more influenced by siblings of the same sex. In Table 3, panel B, we present results separately by whether the

**TABLE 3** Regression results for subgroup analysis, all waves pooled (Waves II, III, and IV), full siblings only

<b>A. When ego is old vs. when ego is young</b>				
	(1) Obese	(2) BMI	(3) Obese	(4) BMI
Mean of DV	Ego = older 27.84%		Ego = younger 25.33%	
Alter's PGS for BMI	-0.0209 [0.0197]	-0.5555 [0.3539]	0.0018 [0.0190]	-0.0043 [0.3274]
Ego's PGS for BMI	0.0728*** [0.0195]	1.1894*** [0.3304]	0.0721*** [0.0192]	1.3432*** [0.3310]
Age difference (in months)	-0.0020 [0.0016]	-0.0212 [0.0290]	-0.0008 [0.0015]	-0.0246 [0.0252]
Same-sex sibling	0.0515 [0.0362]	0.9606 [0.6066]	0.0319 [0.0317]	0.8469* [0.5120]
Observations	1,275	1,275	1,271	1,271
R <sup>2</sup>	.133	.224	.167	.270
<b>B. Same-sex sibling vs. different-sex sibling</b>				
	(1) Obese	(2) BMI	(3) Obese	(4) BMI
Mean of DV	Same-sex siblings 28.34%		Different-sex siblings 24.27%	
Alter's PGS for BMI	-0.0066 [0.0161]	-0.1856 [0.2868]	0.0051 [0.0200]	-0.1711 [0.3075]
Ego's PGS for BMI	0.0934*** [0.0159]	1.6747*** [0.2690]	0.0398* [0.0209]	0.7401** [0.3084]
Age difference (in months)	-0.0032** [0.0016]	-0.0550** [0.0278]	0.0021 [0.0018]	0.0401 [0.0312]
Observations	1,454	1,454	1,092	1,092
R <sup>2</sup>	.170	.266	.135	.222
<b>C. When ego = female vs. when ego = male</b>				
	(1) Obese	(2) BMI	(3) Obese	(4) BMI
Mean of DV	Ego = female 27.29%		Ego = male 25.85%	
Alter's PGS for BMI	-0.0075 [0.0187]	-0.3218 [0.3306]	-0.0052 [0.0172]	-0.1744 [0.2722]
Ego's PGS for BMI	0.0751*** [0.0183]	1.3679*** [0.2982]	0.0705*** [0.0181]	1.1872*** [0.2900]
Age difference (in months)	-0.0014 [0.0015]	-0.0131 [0.0271]	-0.0007 [0.0017]	-0.0196 [0.0282]
Same-sex sibling	0.0068 [0.0357]	0.5063 [0.6161]	0.0322 [0.0373]	0.5200 [0.5973]
Observations	1,312	1,312	1,234	1,234
R <sup>2</sup>	.162	.256	.152	.256
<b>D. When ego = non-white vs. when ego = white</b>				
	(1) Obese	(2) BMI	(3) Obese	(4) BMI
Mean of DV	Ego = non-white 29.21%		Ego = white 25.28%	
Alter's PGS for BMI	-0.0317 [0.0217]	-0.7003* [0.4183]	0.0034 [0.0148]	-0.0475 [0.2288]
Ego's PGS for BMI	0.0654*** [0.0210]	1.1136*** [0.3295]	0.0785*** [0.0155]	1.3891*** [0.2528]
Age difference (in months)	-0.0023 [0.0021]	-0.0383 [0.0365]	-0.0012 [0.0014]	-0.0200 [0.0244]
Same-sex sibling	0.0913* [0.0502]	2.0519** [0.8673]	0.0023 [0.0314]	0.0186 [0.5328]
Observations	849	849	1,697	1,697
R <sup>2</sup>	.200	.293	.136	.231
<b>E. When ego's PGS for BMI &gt; average vs. ego's PGS for BMI &lt; average</b>				
	(1) Obese	(2) BMI	(3) Obese	(4) BMI
Mean of DV	Ego's PGS > average 32.65%		Ego's PGS < average 20.94%	
Alter's PGS for BMI	-0.0131 [0.0221]	-0.4820 [0.4208]	0.0063 [0.0169]	0.0705 [0.2861]
Ego's PGS for BMI	0.0881*** [0.0310]	1.4335*** [0.4893]	0.0253 [0.0246]	0.9677** [0.4072]
Age difference (in months)	-0.0032* [0.0017]	-0.0525* [0.0299]	-0.0002 [0.0015]	-0.0013 [0.0258]
Same-sex sibling	0.0933** [0.0388]	1.9376*** [0.6611]	-0.0041 [0.0333]	-0.0184 [0.5516]
Observations	1,228	1,228	1,318	1,318
R <sup>2</sup>	.143	.231	.148	.246

*Note.* Robust standard errors are in brackets, clustered at the family level. First 10 principal components are included in the model. Demographic variables controlled for are age, sex, race, whether the respondent lives with parents, whether the siblings live together, total household income, the respondent's highest grade completed, marital status, employment status, and parental educational attainment. For the outcome of obese, the ordinary least squares (linear probability model) coefficient is reported.

Abbreviation: DV, dependent variable; PGS, polygenic score.

\*\*\* $p < .01$ . \*\* $p < .05$ . \* $p < .1$ .



siblings are of the same sex or different sex. Again, we see the consistent pattern that the PGS for BMI of the alter has no detectable effect on ego BMI or obesity, irrespective of whether the alter is of the same or different sex than the ego.

Another possibility is that only one of the two sexes is influenced by alters. We investigate this in Table 3, panel C, in which we report results separately for egos who are female and egos who are male. Consistent with the earlier results, we find that alter PGS for BMI is essentially uncorrelated with ego BMI and obesity no matter what the sex of the ego.

As stated in Section 3, the PGS for BMI is based on samples of European ancestry (Locke, Kahali, Berndt, et al., 2015), and as a result, it may be more predictive of BMI for those of European ancestry than for those of other ancestries (see Braudt, 2018; Martin et al., 2017). To investigate this possibility, we estimate our models separately for whites and non-whites in Table 3, panel D. In the model results for non-whites (Table 3, panel D), alter PGS is negatively and significantly correlated with ego BMI (column 2), but we cannot reject the null of no correlation between alter PGS and ego obesity (column 1). The former is the only instance in this study in which we find a statistically significant correlation of alter PGS with ego weight. Taken in the context of all the results, however, we interpret the result as spurious. The significant result is found only for non-whites (not for whites); this is relevant because, as noted earlier, the PGS was developed using white European samples, and thus, it may be less predictive for non-white respondents. In addition, we find the significant result only for the outcome of ego BMI, not for ego obesity. Moreover, the sign of the coefficient is the opposite of what is expected. In addition, as one estimates the base model for a variety of subgroups and robustness checks, it becomes increasingly likely that one will find a statistically significant result by chance (Rothman, 1990). In the context of all these facts, we interpret the one statistically significant finding as likely spurious.

Finally, we examine whether the results differ for egos with an above-average value of the PGS for BMI versus those with a below-average value of the PGS for BMI. The results, shown in Table 3, panel E, indicate that for both groups, the alter PGS for BMI has no detectable effect on ego BMI or obesity.

### 4.3 | Extensions and robustness checks

BMI is an imperfect measure of fatness because it is simply a measure of weight for height; it does not measure body composition (see, e.g., Burkhauser & Cawley, 2008). The Add Health, in Wave IV, includes waist circumference, which is a measure of central adiposity (fatness around the midsection) and is associated with higher risk of cardiovascular disease. A limitation is that waist circumference is only available in Wave IV of Add Health, so there is a limited sample size ( $N = 957$ ). We estimate similar models as before but replace the dependent variable with waist circumference and replace the PGS for BMI with the PGS for waist circumference (which was also established from a GWAS). The pattern of results is the same, however (see Table A2): the ego's own PGS is strongly associated with waist circumference, but the alter's PGS for waist circumference has no detectable effect on the ego's waist circumference, and the point estimates are quite small.

As described in Section 3, we exclude Wave I from our main analysis because in that wave, weight and height were self-reported. However, one advantage to adding Wave I to our sample is that it would increase the sample size and thus provide more precise estimates. Table A3 contains results from models estimated using all four waves of Add Health pooled. The indicator variable for Wave I also controls for the fact that weight and height in that wave are self-reported. The pattern of results is the same as before: For all four waves pooled (Table A3, panel A) and for Wave I separately (panel B), alter's PGS for BMI is essentially uncorrelated with ego's BMI or obesity.

To further address the fact that the PGS is based on a GWAS of people of European descent, we estimate our models using only non-Hispanic whites; this approach is recommended as a robustness check for Add Health genetic data by Braudt and Harris (2018). The results, which appear in Table A4, panel A (Waves II, III, and IV pooled) and panels B through D (for each of those waves separately), follow the same pattern as the earlier results—the ego's PGS for BMI is highly correlated with their BMI and obesity, but alter's PGS for BMI has no detectable correlation with ego weight.

One potential concern is that our main model (with results in Table 2) includes as regressors some potentially endogenous variables: income, education, and marital status. As a robustness check, we re-estimated the model controlling for nothing other than the PGS of the ego and the PGS of the alter and then re-estimated it again by adding only controls for the strictly exogenous variables age, sex, and race. The results, which are available upon request, are very similar to our main results; the point estimates are almost identical.

## 5 | DISCUSSION

This paper tests for a form of peer effects—genetic nurture—using novel data. Specifically, we test whether BMI and the probability of obesity are associated with a sibling's PGS for BMI, controlling for the individual's own PGS. The results

are extremely consistent: We cannot reject the null hypothesis of no genetic nurture. The coefficients on alter PGS are not only statistically insignificant, but they are also quite small in magnitude. This is true for a wide variety of samples: For the three waves with measured BMI, for all four waves (including the one in which BMI is based on self-reports), for each wave individually, for older alters and younger alters, for alters of the same sex and different sex as the ego, for male and female egos, and for egos with above-average and below-average PGS. We also find the same pattern of results when we estimate models for waist circumference, although that information is only available in a single wave, and thus, it has a small sample size. In one case, we did find a statistically significant correlation between alter PGS and ego BMI, among non-whites, but we interpret this as spurious because the PGS was developed from a white sample and is stated to be less predictive for non-whites, the coefficient is of the opposite sign one would predict, and the result is not robust to using obesity rather than BMI as a measure of weight. To summarize the entire set of results, we find no credible evidence that there is genetic nurture in BMI or obesity between siblings.

This evidence is consistent with two previous studies that used different methods and samples from other countries and examined a different family relationship (parents to children). Kong et al. (2018) and Richmond et al. (2017) examined the association of parental alleles for BMI that are not transmitted to children, with the child's BMI. They too find no evidence of genetic nurture in BMI, in their case between parents and children. The results of these studies, which are based on data from Iceland and the United Kingdom, respectively, are consistent with and complement the results of this study based on U.S. data. It is also noteworthy that Kong et al. (2018) did find evidence of genetic nurture with respect to other outcomes. They found that the effect of nontransmitted alleles associated with education had 29.9% of the effect of transmitted alleles. They also found effects on age at first child, high-density lipoprotein level, fasting glucose level, height, and cigarettes smoked per day. Rauscher, Conley, and Siegal (2015) found evidence of genetic nurture among siblings in the Add Health data but for health-related absences and self-rated health. These suggest that our null finding for BMI is not simply an artifact of the general approach.

Our inability to reject the null of no genetic nurture in BMI is consistent with a large literature that found no detectable effect of shared household environment on BMI or obesity in adolescents and adults (Haberstick et al., 2010; Nan, Guo, Warner, et al., 2012; Silventoinen, Rokholm, Kaprio, & Sorensen, 2010).<sup>11</sup>

The results of this paper are useful for interpreting other research in health economics. Numerous studies have documented a substantial correlation in sibling BMI (e.g., Brown & Roberts, 2012); the results of this paper cast doubt on whether that similarity is due to genetic nurture within families. Instead, the similarity is more likely due to other factors such as similarity of genetic endowments, parental reallocations of resources across siblings to achieve equality of outcomes, or parental learning by doing in production of their children's health (e.g., Griliches, 1979; Yi, Heckman, et al. 2015).

This paper also relates to the literature that found evidence of peer effects in weight (Yakusheva, Kapinos, & Eisenberg, 2014) and in fitness (Carrell, Hoekstra, & West, 2011) among a different set of peers—fellow college students. This paper also contributes to the broader literature on sibling spillovers in health and health-care utilization, for example, Altonji, Cattan, and Ware (2017), Breining, Daysal, Simonsen, and Trandafir (2015), Adhvaryu and Nyshadham (2016), and Ho (2017).

When interpreting these results, the following caveats should be kept in mind. First, we study only one form of genetic nurture: that among full siblings. This is advantageous because it is an exogenously determined peer group; any similarity cannot be due to self-selection, which is a factor that complicates study of other peer groups such as friends or neighbors. However, we acknowledge that other peer groups may exhibit different effects of genetic nurture. The fact that we cannot reject the null of no genetic nurture in BMI does not mean that there may not be genetic nurture with respect to other characteristics (although we also find no evidence of genetic nurture for waist circumference).

Other limitations of the study include that although the PGS is the most recently available measure of genetic endowment relating to BMI and has nearly double the explanatory power of the prior measure, it still only explains 2.7% of the variance in BMI, which is far less than the total variance in BMI that is believed to be genetic. In other words, it is believed that much of the genetic variation in BMI has yet to be linked to specific SNPs. The scientific frontier is always progressing, and this study should be replicated when more powerful PGS become available.<sup>12</sup> Likewise, when GWAS is conducted for African-American and/or Hispanic populations, the PGS developed for those groups should be used to test for genetic nurture in those populations.

<sup>11</sup>Research based on twin and adoption studies has found that any common environment effects in BMI disappear by early adolescence (e.g., Nan, Guo, Warner, et al., 2012; Silventoinen, Rokholm, Kaprio, & Sorensen, 2010), and some studies find no common environment effects in BMI at any age (Haberstick et al., 2010).

<sup>12</sup>For example, a broader PGS based on a new GWAS reportedly explains about 6% of the variation in BMI for people of European ancestry (Yengo et al., 2018). However, this PGS is not available in the Add Health.

Another issue for generalizability is that we examine siblings during the ages of 12 to 34 years. Although we cannot reject the null hypothesis of no genetic nurture among siblings in these ages, it is possible that some could exist at earlier or later ages. However, it should be noted that substantial correlations in numerous health behaviors and health conditions exist at the ages we examine (e.g., Daw, Margolis, & Verdery, 2015), so we are examining a section of the life course in which siblings exhibit similar behaviors and traits.

The null results in this paper have implications for policy. We find no evidence of social contagion in BMI or obesity. This is useful because, as mean BMI worldwide rose between 1975 and 2014 from 21.7 to 24.2 for men and from 22.1 to 24.4 for women (NCD Risk Factor Collaboration, 2016), there has been concern that peer effects may lead rising BMI to become self-perpetuating. Because we fail to reject the null of no genetic nurture among siblings and find very small point estimates, this study yields no evidence that such spillovers exist, at least among siblings. The downside of this null result is that it also implies that successful weight reduction strategies may not have beneficial spillover effects within families, which, if they existed, would increase their cost-effectiveness.

## CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.


## ACKNOWLEDGEMENTS

For helpful feedback, we thank Sonja Kassenboehmer, Lauren Schmitz, Olga Yakusheva, Dalton Conley, Justin Trogdon, Pietro Biroli, Hans van Kippersluis, and participants at the European Workshop on Econometrics and Health Economics, the NBER Health Economics meetings, and the kickoff meeting of the NORFACE project “Gene-Environment Interplay in the Generation of Education and Health Inequalities” (Grant 462-16-100) at the Erasmus School of Economics. This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is for Ronald R. Rindfuss and Barbara Entwisle for their assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from Grant P01-HD31921 for this analysis.

## ORCID

John Cawley  <https://orcid.org/0000-0002-4805-9883>

Euna Han  <https://orcid.org/0000-0003-2656-7059>

Edward C. Norton  <https://orcid.org/0000-0003-4555-0631>

## REFERENCES

- Adhvaryu, A., & Nyshadham, A. (2016). Endowments at birth and parents' investments in children. *Economic Journal*, 126(593), 781–820.
- Almond, D., Edlund, L., & Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124(4), 1729–1772.
- Altonji, J. G., Cattan, S., & Ware, I. (2017). Identifying sibling influence on teenage substance use. *Journal of Human Resources*, 52(1), 1–47.
- Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., ... Simonsen, M. (2017). Sibling spillovers. NBER Working Paper #23062.
- Braudt, D. B. (2018). Polygenic scores in Add Health. Retrieved from [https://www.cpc.unc.edu/projects/addhealth/events/Add\\_Health\\_Genetic\\_Data\\_and\\_GenomeWide\\_Association\\_Study\\_Braudt.pdf](https://www.cpc.unc.edu/projects/addhealth/events/Add_Health_Genetic_Data_and_GenomeWide_Association_Study_Braudt.pdf)
- Braudt, D. B., & Harris, K. M. (2018). Polygenic scores (PGSs) in the National Longitudinal Study of Adolescent to Adult Health (Add Health) —Release 1. Retrieved from <https://cdr.lib.unc.edu/record/uuid:98614f6f-84ce-43b6-b899-5c105916eae>
- Breining, S., Daysal, N. M., Simonsen, M., & Trandafir, M. (2015). Spillover effects of early-life medical interventions. IZA DP #9086.
- Breining, S. N. (2014). The presence of ADHD: Spillovers between siblings. *Economics Letters*, 124, 469–473.

- Brown, H. W., & Roberts, J. (2012). Exploring the factors contributing to sibling correlations in BMI: A study using the panel study of income dynamics. *Obesity, 20*, 978–984. <https://doi.org/10.1038/oby.2011.351>
- Brownell, K. D., & Horgen, K. B. (2004). *Food fight: The inside story of the food industry, America's obesity crisis, and what we can do about it*. New York: McGraw Hill.
- Burkhauser, R. V., & Cawley, J. (2008). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics, 27*(2), 519–529. <https://doi.org/10.1016/j.jhealeco.2007.05.005>
- Carrell, S. E., Hoekstra, M., & West, J. E. (2011). Is poor fitness contagious? Evidence from randomly assigned friends. *Journal of Public Economics, 95*(7–8), 657–663.
- Cawley, J. (2015). An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. *Journal of Health Economics, 43*, 244–268. <https://doi.org/10.1016/j.jhealeco.2015.03.001>
- Cawley, J., Maclean, J. C., Hammer, M., & Wintfeld, N. (2015). Reporting error in weight and its implications for bias in economic models. *Economics and Human Biology, 19*, 27–44. <https://doi.org/10.1016/j.ehb.2015.07.001>
- Cohen-Cole, E., & Fletcher, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics, 27*(5), 1382–1387. <https://doi.org/10.1016/j.jhealeco.2008.04.005>
- Conley, D., Laidley, T. M., Boardman, J. D., & Domingue, B. W. (2016). Changing polygenic penetrance on phenotypes in the 20th century among adults in the US population. *Scientific Reports, 6*, 30348. <https://doi.org/10.1038/srep30348>
- Corman, H., Dave, D. M., & Reichman, N. E. (2017). Evolution of the infant health production function. NBER Working Paper #24131.
- Daw, J., Margolis, R., & Verdery, A. M. (2015). Siblings, friends, course-mates, club-mates: How adolescent health behavior homophily varies by race, class, gender, and health status. *Social Science and Medicine, 125*, 32–39. <https://doi.org/10.1016/j.socscimed.2014.02.047>
- Domingue, B. W., Belsky, D. W., Fletcher, J. M., Conley, D., Boardman, J. D., & Harris, K. M. (2018). The social genome of friends and school-mates in the National Longitudinal Study of Adolescent to Adult Health. *Proceedings of the National Academy of Sciences of the United States of America, 115*(4), 702–707. <https://doi.org/10.1073/pnas.1711803115>
- Fletcher, J., Hair, N. L., & Wolfe, B. (2012). Am I my brother's keeper? Sibling spillover effects: The case of developmental disabilities and externalizing behavior. NBER working paper #18279.
- Fletcher, J., Mailick, M., Song, J., & Wolfe, B. (2013). A sibling death in the family: Common and consequential. *Demography, 50*, 803–826.
- Fletcher, J., Vidal-Fernandez, M., & Wolfe, B. (2018). Dynamic and heterogeneous effects of sibling death on children's outcomes. *Proceedings of the National Academy of Sciences of the United States of America, 115*(1), 115–120. <https://doi.org/10.1073/pnas.1709092115>
- Fletcher, J., & Wolfe, B. (2008). Child mental health and human capital accumulation: The case of ADHD revisited. *Journal of Health Economics, 27*, 794–800. <https://doi.org/10.1016/j.jhealeco.2007.10.010>
- Griliches, Z. (1979). Sibling models and data in economics—Beginnings of a survey. *Journal of Political Economy, 87*(5), S37–S64.
- Guo, G., Liu, H., Wang, L., Shen, H., & Hu, W. (2015). The genome-wide influence on human BMI depends on physical activity, life course, and historical period. *Demography, 52*(5), 1651–1670. <https://doi.org/10.1007/s13524-015-0421-2>
- Haberstick, B. C., Lessem, J. M., McQueen, M. B., Boardman, J. D., Hopfer, C. J., Smolen, A., & Hewitt, J. K. (2010). Stable genes and changing environments: Body mass index across adolescence and young adulthood. *Behavior Genetics, 40*(4), 495–504.
- Hales, C. M., Fryar, C. D., Carroll, M. D., Freedman, D. S., & Ogden, C. L. (2018). Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007–2008 to 2015–2016. *JAMA, 319*(16), 1723–1725. <https://doi.org/10.1001/jama.2018.3060>
- Halliday, T. J., & Mazumder, B. (2017). An analysis of sibling correlations in health using latent variable models. *Health Economics, 26*(12), E108–E125. <https://doi.org/10.1002/hec.3483>
- Ho, C. Y. (2017). Estimating sibling spillovers in health: Evidence on symptoms. *Economics and Human Biology, 27*, 93–101. <https://doi.org/10.1016/j.ehb.2017.05.003>
- Hu, F. (2008). *Obesity epidemiology*. New York: Oxford University Press.
- Institute of Medicine (2012). *Accelerating progress in obesity prevention: Solving the weight of the nation*. Washington, DC: The National Academies Press.
- Kong, A., Thorleifsson, G., Frigge, M. L., Michael, L., Vilhjalmsdottir, B. J., Young, A. I., ... Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science, 359*(6374), 424–428. <https://doi.org/10.1126/science.aan6877>
- Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *Quarterly Journal of Economics, 64*(2), 183–207.
- Liu, H., & Guo, G. (2015). Lifetime socioeconomic status, historical context, and genetic inheritance in shaping body mass in middle and late adulthood. *American Sociological Review, 80*(4), 705–737. <https://doi.org/10.1177/0003122415590627>
- Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., ... Croteau-Chonka, D. C. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature, 518*(7538), 197–206. <https://doi.org/10.1038/nature14177>
- Manski, C. F. (1993). Identification of endogenous social effects—The reflection problem. *Review of Economic Studies, 60*(3), 531–542.
- Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives, 14*(3), 115–136.

- Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., ... Kenny, E. E. (2017). Human demographic history impacts genetic risk prediction across diverse populations. *American Journal of Human Genetics*, *100*(4), 635–649. <https://doi.org/10.1016/j.ajhg.2017.03.004>
- McQueen, M. B., Boardman, J. D., Domingue, B. W., Smolen, A., Tabor, J., Killeya-Jones, L., ... Harris, K. M. (2015). The National Longitudinal Study of Adolescent to Adult Health (Add Health) sibling pairs genome-wide data. *Behavior Genetics*, *45*(1), 12–23. <https://doi.org/10.1007/s10519-014-9692-4>
- Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the United States, 2000. *JAMA*, *291*, 1238–1245. <https://doi.org/10.1001/jama.291.10.1238>
- Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2005). Correction: Actual causes of death in the United States, 2000. *JAMA*, *293*, 293–294. <https://doi.org/10.1001/jama.293.3.293>
- Nan, C., Guo, B., Warner, C., Fowler, T., Barrett, T., Boomsma, D., ... Zeegers, M. (2012). Heritability of body mass index in pre-adolescence, young adulthood and late adulthood. *European Journal of Epidemiology*, *27*, 247–253. <https://doi.org/10.1007/s10654-012-9678-6>
- NCD Risk Factor Collaboration (NCD-RisC) (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet*, *387*, 1377–1396.
- Nestle, M. (2002). *Food politics: How the food industry influences nutrition and health*. Los Angeles: University of California Press.
- Parman, J. (2015). Childhood health and sibling outcomes: Nurture reinforcing nature during the 1918 influenza pandemic. *Explorations in Economic History*, *58*, 22–43.
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. <https://doi.org/10.1038/ng1847>
- Rauscher, E., Conley, D., & Siegal, M. L. (2015). Sibling genes as environment: Sibling dopamine genotypes and adolescent health support frequency dependent selection. *Social Science Research*, *54*, 209–220. <https://doi.org/10.1016/j.ssresearch.2015.08.002>
- Richmond, R. C., Timpson, N. J., Felix, J. F., Palmer, T., Gaillard, R., McMahon, G., ... Lawlor, D. A. (2017). Using genetic variation to explore the causal effect of maternal pregnancy adiposity on future offspring adiposity: A Mendelian randomisation study. *PLoS Medicine*, *14*(1), 1002221. <https://doi.org/10.1371/journal.pmed.1002221>
- Rosenquist, J. N., Lehrer, S. F., James O'Malley, A., Zaslavsky, A. M., Smoller, J. W., & Christakis, N. A. (2015). Birth cohort modifies FTO-BMI association. *Proceedings of the National Academy of Sciences*, *112*(2), 354–359. <https://doi.org/10.1073/pnas.1411893111>
- Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, *1*(1), 43–46.
- Sassi, F. (2010). *Obesity and the economics of prevention: Fit not fat*. Paris, France: OECD. [https://58.27.242.36:8000/jspui/bitstream/1/10/1/Fit\\_not\\_Fat\\_Obesity\\_and\\_the\\_Economics\\_of\\_Prevention.PDF](https://58.27.242.36:8000/jspui/bitstream/1/10/1/Fit_not_Fat_Obesity_and_the_Economics_of_Prevention.PDF)
- Schmitz, L. L., & Conley, D. (2018). The impact of late-career job loss and genotype on body mass index. NBER Working Paper #22348.
- Silventoinen, K., Rokholm, B., Kaprio, J., & Sorensen, T. I. (2010). The genetic and environmental influences on childhood obesity: A systematic review of twin and adoption studies. *International Journal of Obesity*, *34*, 29–40. <https://doi.org/10.1038/ijo.2009.177>
- Smolen, A., Whitsel, E. A., Tabor, J., Killeya-Jones, L. A., Cuthbertson, C. C., Hussey, J. M., ... Harris, K. M. (2013). Add Health wave IV documentation: Candidate genes. Retrieved from [https://www.cpc.unc.edu/projects/addhealth/documentation/guides/DNA\\_documentation.pdf](https://www.cpc.unc.edu/projects/addhealth/documentation/guides/DNA_documentation.pdf)
- Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., ... Loos, R. J. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics*, *42*, 937–948. <https://doi.org/10.1038/ng.686>
- Ware, E. B., Schmitz, L. L., Faul, J., Gard, A., Mitchell, C., Smith, J. A., ... Kardia, S. LR. (2017). Heterogeneity in Polygenic Scores for Common Human Traits. *BioRxiv*, 106062.
- Wolf, J. B., Brodie, E. D. III, Cheverud, J. M., Moore, A. J., & Wade, M. J. (1998). Evolutionary consequences of indirect genetic effects. *Trends in Ecology & Evolution*, *13*(2), 64–69.
- Yakusheva, O., Kapinos, K. A., & Eisenberg, D. (2014). Estimating heterogeneous and hierarchical peer effects on body weight using roommate assignments as a natural experiment. *Journal of Human Resources*, *49*(1), 234–261.
- Yengo, L., Sidorenko, J., Kemper, K. E., Zheng, Z., Wood, A. R., Weedon, M. N., ... Consortium, G. I. A. N. T. (2018). Meta-analysis of genome-wide association studies for height and body mass index in ~700 000 individuals of European ancestry. *Human Molecular Genetics*, *27*(20), 3641–3649. <https://doi.org/10.1093/hmg/ddy271>
- Yi, J. J., Heckman, J. J., Zhang, J. S., & Conti, G. (2015). Early health shocks, intra-household resource allocation and child outcome. *Economic Journal*, *125*(588), F347–F371.

**How to cite this article:** Cawley J, Han E, Kim J, Norton EC. Testing for family influences on obesity: The role of genetic nurture. *Health Economics*. 2019;28:937–952. <https://doi.org/10.1002/hec.3889>

## APPENDIX

**TABLE A1** Balance table comparing respondent and household characteristics by whether respondent has PGS for BMI or not (Add Health, Wave II)

Variable Name	PGS reported ( <i>n</i> = 1,659)	PGS missing ( <i>n</i> = 3,325)	Difference	<i>T</i> statistics
Respondents' characteristics				
Boy	0.48	0.49	−0.01	−0.65
Age	16.43	16.52	−0.09*	−1.76
White	0.64	0.45	0.19***	13.48
African-American	0.21	0.25	−0.05***	3.64
Hispanic	0.08	0.18	−0.10***	−10.49
Other race	0.06	0.11	−0.05***	−6.28
Live with parents	0.90	0.88	0.02*	1.86
Height (in.)	66.16	66.04	0.12	0.86
Weight (lb)	140.95	139.34	1.61	1.43
BMI	23.11	22.86	0.25*	1.65
Obese	0.14	0.13	0.01	1.32
Household characteristics				
20K < income < 39K	0.23	0.23	0.00	0.32
40K < income < 49K	0.09	0.09	0.00	0.02
50K < income < 75K	0.19	0.16	0.03**	2.74
75K < income < 99K	0.06	0.05	0.01**	2.09
100K < income	0.05	0.04	0.01**	2.43
Parents married	0.63	0.59	0.03**	2.21
Parents: less than HS	0.12	0.18	−0.06***	−5.67
Parents: HS grad	0.28	0.25	0.05***	3.98
Parents: less than college	0.27	0.25	0.03*	1.95
Parents: college grad	0.19	0.18	0.01	0.92
Parents: employed	0.51	0.48	0.02	1.63

Abbreviations: HS, high school; PGS, polygenic score.

**TABLE A2** Regression results using PGS for waist circumference, full siblings only (Wave IV)

	(1)	(2)
	Waist circumference (cm)	
Mean of DV	98.98	
Alter's PGS for waist circumference		−0.2469 [0.6317]
Ego's PGS for waist circumference	2.5234*** [0.5490]	2.6559*** [0.5881]
Observations	957	957
<i>R</i> <sup>2</sup>	.093	.094

*Note.* Robust standard errors are in brackets, clustered at the family level. First 10 principal components are included in the model. Demographic variables controlled for are age, sex, race, whether the respondent lives with parents, whether the siblings live together, total household income, the respondent's highest grade completed, marital status, employment status, and parental educational attainment. For the outcome of obese, the ordinary least squares (linear probability model) coefficient is reported.

Abbreviation: DV, dependent variable; PGS, polygenic score.

\*\*\**p* < .01. \*\**p* < .05. \**p* < .1.

**TABLE A3** Regression results for Equation (1), full siblings only (Wave I included)

<b>A. All waves pooled (Waves I, II, III, and IV)</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	22.84%		25.67	
Alter's PGS for BMI		-0.0088 [0.0116]		-0.2701 [0.1953]
Ego's PGS for BMI	0.0572*** [0.0115]	0.0616*** [0.0115]	0.9618*** [0.1920]	1.0993*** [0.1824]
Observations	3,480	3,480	3,480	3,480
R <sup>2</sup>	.117	.117	.235	.236
<b>B. Wave I</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	12.63%		22.73	
Alter's PGS for BMI		-0.0091 [0.0126]		-0.2235 [0.1683]
Ego's PGS for BMI	0.0374*** [0.0120]	0.0420*** [0.0130]	0.6790*** [0.1650]	0.7916*** [0.1615]
Observations	934	934	934	934
R <sup>2</sup>	.048	.048	.080	.081

Note. Robust standard errors are in brackets, clustered at the family level. First 10 principal components are included in the model. Demographic variables controlled for are age, sex, race, whether the respondent lives with parents, whether the siblings live together, total household income, the respondent's highest grade completed, marital status, employment status, and parental educational attainment. For the outcome of obese, the ordinary least squares (linear probability model) coefficient is reported.

Abbreviation: DV, dependent variable; PGS, polygenic score.

\*\*\*p < .01. \*\*p < .05. \*p < .1.

**TABLE A4** Regression results for Equation (1), non-Hispanic white full siblings only

<b>A. All waves pooled (Waves II, III, and IV)</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	25.39%		26.58	
Alter's PGS for BMI		0.0031 [0.0149]		-0.0540 [0.2294]
Ego's PGS for BMI	0.0793*** [0.0160]	0.0777*** [0.0157]	1.3551*** [0.2629]	1.3823*** [0.2563]
Observations	1,689	1,689	1,689	1,689
R <sup>2</sup>	.121	.121	.214	.214
<b>B. Wave II</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	15.45%		23.43	
Alter's PGS for BMI		-0.0138 [0.0173]		-0.0565 [0.2195]
Ego's PGS for BMI	0.0755*** [0.0184]	0.0826*** [0.0191]	1.2764*** [0.2697]	1.3055*** [0.2593]
Observations	563	563	563	563
R <sup>2</sup>	.084	.085	.119	.119
<b>C. Wave III</b>				
	(1)	(2)	(3)	(4)
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	23.54%		26.87	
Alter's PGS for BMI		0.0069 [0.0210]		-0.0652 [0.2773]
Ego's PGS for BMI	0.0607*** [0.0220]	0.0574** [0.0233]	1.1439*** [0.3066]	1.1749*** [0.3040]
Observations	497	497	497	497
R <sup>2</sup>	.118	.119	.117	.117

<b>D. Wave IV</b>				
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>
	<b>Obese</b>		<b>BMI</b>	
Mean of DV	35.77%		29.18	
Alter's PGS for BMI		0.0202 [0.0206]		0.0479 [0.3067]
Ego's PGS for BMI	0.0816*** [0.0210]	0.0712*** [0.0220]	1.3768*** [0.3254]	1.3519*** [0.3339]
Observations	629	629	629	629
R <sup>2</sup>	.116	.117	.128	.128

*Note.* Robust standard errors are in brackets, clustered at the family level. First 10 principal components are included in the model. Demographic variables controlled for are age, sex, whether the respondent lives with parents, whether the siblings live together, total household income, the respondent's highest grade completed, marital status, employment status, and parental educational attainment. For the outcome of obese, the ordinary least squares (linear probability model) coefficient is reported.

Abbreviation: DV, dependent variable; PGS, polygenic score.

\*\*\* $p < .01$ . \*\* $p < .05$ . \* $p < .1$ .