Published in final edited form as:

Econ Hum Biol. 2018 September ; 31: 200–208. doi:10.1016/j.ehb.2018.08.011.

Does Education Attenuate the Genetic Risk of Obesity? Evidence from U.K. Twins.

Vikesh Amin,

Department of Economics, Central Michigan University, Mt Pleasant, MI 48859. Tel:
989-774-3373. Fax 989-774-2040

Paul Dunn, and

Department of Business Information Systems, Central Michigan University, Mt Pleasant, MI 4885.

Tim Spector

Department of Twin Research & Genetic Epidemiology, Kings College London, Strand, London,
WC2R 2LS, UK.

Abstract

More education is associated with a lower body mass index (BMI) and likelihood of being obese. Obesity and BMI also have a strong genetic basis. Given these observations, we investigate whether more education can reduce obesity by attenuating the underlying genetic risk of being obese, through gene-environment (GxE) interactions. We estimate associations between (i) education, (ii) a genetic risk score (GRS) and (iii) GxE interactions between education and the GRS through Ordinary least Squares (OLS) and twins fixed-effect regressions using the TwinsUK database. OLS estimates show that there are significant associations of education and genetics. Twins with a university education are 14.3 percentage points less likely to be obese compared to twins with less than compulsory education, and a 1 standard deviation increase in the GRS increases the likelihood of being obese by 5.2 percentage points. The GxE interactions are statistically insignificant, suggesting that the marginal association of the GRS with obesity does not differ by educational attainment. When controlling for confounding through twins fixed-effects, we find a smaller role of genetics. The association of the GRS with obesity decreases to 0.040. Associations of educational attainment are substantially reduced and insignificant. GxE interactions also remain insignificant. Overall, we find little evidence of any GxE interactions.

Keywords

Education; Obesity; Gene-Environment Interactions; Twins

(Corresponding Author): Department of Economics, Central Michigan University, Mt Pleasant, MI 48859. amin1v@cmich.edu. Tel: 989-774-3373. Fax 989-774-2040. dunn1pt@cmich.edu, tim.spector@kcl.ac.uk.

⁹We exclude twins who are under 25 years of age when BMI was measured because education may not have been completed

1. Introduction

A large empirical literature has established that more education is associated with a lower body mass Index (BMI) and a lower likelihood of being obese¹. Education can affect obesity and BMI through several channels. More education can reduce the likelihood of being obese because education increases resources enabling purchase of better health care (Chandola et al. 2006); increases access to and use of information in making health-related decisions (Kenkle 1991); improves one's rank in society, which is associated with better adult health and reduced stress (Rose and Marmot 1981); increases interactions with other more-educated peers generating health spillovers (Fletcher 2010); and changes preferences, norms, and behaviors related to health (Becker and Mulligan 1997).

We contribute to the education-health literature by testing a different hypothesis: does more education reduce the underlying genetic risk of being obese and having a high BMI? Obesity and BMI have a strong genetic basis. Using behavioral genetics twin-based models Maes et al. (1997) find that 40%–70% of the variation in BMI is due to genetics. Based on an adoption design, Classen and Thompson (2016) find that child BMI/obesity is uncorrelated with the BMI/obesity of adoptive parents, indicating that the intergenerational transmission of obesity is due to genetics. Given the importance of genetics, more education may reduce obesity and BMI by attenuating the underlying genetic risk through gene-environment (GxE) interactions. GxE interactions occur when the effect of genetic risk varies across the environment (in this case education). There are theoretical models suggesting that education may reduce the genetic risk. The diathesis-stress model (Ellis et al. 2011) hypothesizes that unhealthy environments (low education) trigger risk alleles, while healthy environments (high education) protect against risk alleles. A college graduate with a high genetic risk of being obese may never be obese because education is associated with higher income affording the individual to eat healthier and having peers who are conscientious about leading a healthy lifestyle. In contrast, a high school dropout with a high genetic risk is more likely to be obese because he/she has a low income, is more likely to consume fast food, and has peers who are also obese. The differential susceptibility (also commonly known as the orchids & dandelions) model (Belsky & Pluess 2009) hypothesizes that some individuals are dandelions, meaning they have a genetic make-up that is unaffected by both positive and negative environments. Others are orchids whose genetic make-up is highly sensitive to the environment. These individuals thrive in positive environments but wilt in negative environments. In economics, Biroli (2015) incorporates genetic risk into Grossman's (1972) model of health production. In the model genes interact with the environment by changing the productivity and implicit costs of health investments. The productivity effect means that for the same level of investments, an individual with a higher genetic risk of obesity is more likely to be obese, than an individual with a lower genetic risk. The cost effect means an individual who is genetically predisposed to obesity may also find it harder and costlier to

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

engage in health behaviors such as exercise. The productivity and cost effects lead to changes in the optimal level of investments, which in turn affects health².

To investigate this hypothesis we use data on MZ (monozygotic; identical) and DZ (dizygotic; fraternal) twins from the TwinsUK database and estimate GxE regressions. We first treat the twins as unrelated individuals and estimate main associations of (i) education (ii) a genetic risk score (GRS) which is used to measure genetic predisposition, and (iii) GxE interactions between the GRS and education through Ordinary Least Squares (OLS) regressions. However, OLS estimates are unlikely to represent causal relations, because genetic risk and education could be endogenous, leading to biased estimates. Genetic risk could be endogenous because of population stratification, which induces a spurious correlation between the GRS and obesity. As education is not randomly assigned, it is likely to be endogenous because it is correlated with unobserved individual and family characteristics (e.g. ability, health endowments, family background) that also affect obesity. To control for some (but not all) of the biases, we exploit the twin dimension and perform twins fixed-effect regressions. Twins fixed-effects relates within-twin pair differences in BMI/obesity to within-twin pair differences in education and genetics. This controls for population stratification, and all unobserved genetic and environmental factors that are shared by twins since conception. We then compare the OLS and twins fixed-effects estimates to assess the magnitude and importance of confounding. Our main contribution is the use of a GRS *and* twins fixed-effects to control for confounding. The use of a GRS to measure genetic risk is the current best practice in the literature. While previous studies have used GRSs, they have not taken account of confounding. To the best of our knowledge, only one previous study (Cook & Fletcher 2015) used sibling fixed-effects. However, they used a single gene to measure genetic risk, and this approach is rarely used now.

The paper is organized as follows. Section 2 provides a background on issues involved in estimating GxE interactions. In particular, it explains the approaches used to measure genetic risk, and discusses the endogeneity problem. Section 3 reviews previous studies that estimate GxE interactions on health where education is the environment. We explain how twins fixed-effects controls for confounding in section 4. We describe dataset in section 5. The results are presented and discussed in section 6. Finally, section 7 concludes.

2. Background

GxE interactions can be estimated using OLS regression for equation (1) where the health outcome of individual i (Y_i) is related to a measure of genetic risk (G_i), educational attainment (E_i), an interaction between the two ($E_i * G_i$), a vector of control variables (X_i) and an error term (u_i).

$$Y_i = b_0 + b_1 G_i + b_2 E_i + b_3 (E_i * G_i) + X_i' b_4 + u_i \quad (1)$$

¹We refer readers to Cohen et al. (2003) for a review of the literature.

The key parameter is b_3 which measures the GxE interaction. b_3 represents the marginal association of genetic risk for different levels of educational attainment. If there are GxE interactions with respect to education, then b_3 will be negative and statistically significant. We explain below, two key issues involved in estimating GxE interactions.

2.1 Approaches to measure genetic risk

Genetic risk was initially measured using single genetic markers (e.g., the FTO gene which is associated with obesity), mainly because of data limitations. Initial genetic markers that were made available were those that were hypothesized to be the main importance. These markers (called candidate genes) were generally chosen to be genotyped because they were located on a particular chromosome region suspected of being involved in the outcome. However, this approach is rarely used now for several reasons. First, complex health outcomes are highly polygenic and have a continuum of genetic risk that reflects small contributions from many genetic loci (Visscher et al. 2008). Second, genetic predisposition to health is also quantitative. Some individuals carry few risk alleles, while others carry many, and most of the population is somewhere in the middle (Plomin et al. 2009). The use of single genetic markers is unlikely to be informative regarding where an individual is located on the continuum of genetic risk. Third, the use of single genes is often underpowered to detect interactions effects (Conley 2015).

For these reasons, the current best practice in the literature is to use a GRS. GRSs are constructed using results from genome-wide association studies (GWAS). A GWAS is an atheoretical approach to gene discovery where hundreds of thousands of single nucleotide polymorphisms (SNPs) are tested for associations with an outcome³. A GWAS consists of a discovery and a replication phase. In the discovery phase, a GWAS pools large consortia of genetic data using meta-analyses and runs regressions testing each SNP at the genome-wide significance level of 5×10^{-8} . Significant associations found in the discovery phase are then tested in independent samples in the replication phase. A GRS aggregates the SNPs identified in a GWAS and weights them by the strength of their associations to (i) produce a single quantitative measure of genetic risk and (ii) increase power. Specifically, the GRS for individual i is a weighted average across the number of SNPs (n) of the number of reference alleles A (0, 1 or 2) at that SNP multiplied by the corresponding beta estimate from the GWAS analysis.

$$GRS_i = \sum_{j=1}^n \beta_j A_{ij} \quad (2)$$

2.2 Endogeneity issues

The key assumption needed for OLS estimates to represent causal effects is that education, genetic risk, and their interactions are uncorrelated with the error term (i.e. exogenous).

²To test the model, Biroli takes physical activity and food intake as health investments and evaluates their interaction with the FTO gene, which is associated with obesity. He finds that genetic factors change both the production function of BMI and level of health investments. Specifically, adolescents carrying the riskier allele in the FTO genotype are at greater risk of obesity even when consuming the same calories and performing the same physical activities.

However, this assumption is unlikely to hold. Even though genes are randomly inherited from parents at conception, there could be a correlation between genetic risk and the error term because of population stratification. Population stratification refers to a situation where the distribution of genes systematically differ by population subgroups (e.g. by ethnicity/race). If these subpopulations also systematically have different health outcomes that are not due to genetic make-up, then this leads to a spurious correlation between genetic risk and health. Population stratification can be controlled for by limiting analyses to ethnically homogenous samples (Cardon and Palmer 2003), by including principal components from genome-wide SNP data as control variables in equation (1) (Price et al. 2006), or by controlling for family background with family fixed-effects analyses (Conley and Rauscher 2013). Therefore, conditional on population stratification, genetic risk is uncorrelated with the error term. However, education is still likely to be endogenous because it is correlated with unobserved factors such as ability, health endowments, motivation, family background that also affect health. For example, birthweight is often taken as a proxy for innate health and ability. Several studies using within-twin and within-sibling comparisons have found significant effects of birthweight on cognitive development (Figlio et al. 2014), educational attainment (Royer 2009), adult health (Black et al. 2007, Iliadou et al. 2004). A failure to control for unobserved innate ability and health will lead to biased estimates, because these unobserved factors are related to both education and health. Education could also be endogenous because of reverse causality, insofar as early-life health affects educational attainment. Several studies have found that childhood obesity is associated with lower academic achievement (Datar and Strum 2006, Sabia 2007, Kaestner and Grossman 2009). Gene-environment correlations can also lead to education being endogenous. This refers to a situation where genes influence an individual's environment and health. Therefore, unless one uses variation in education which is uncorrelated with the error term, education and its interaction with genetic risk are likely to be endogenous leading to biased estimates. Indeed, in a review paper on molecular genetics for economists Lundborg and Stenberg (2010 pg. 327) note "for any policy an important reservation is that the GxE findings must reflect causal effects".

3. Previous studies

Only a small number of studies that have estimated main effects of education, genetic risk, and GxE interactions on health. Some studies use single genetic markers to measure genetic predisposition. Mitchell et al. (2011) estimate interactions between education and the number of 5-HTTLPR alleles on a mother's risk of postpartum depression using a sample of disadvantaged women from the Fragile Families database. Regression estimates from equation (1) indicate that there are no main effects of education and genetic risk. However, they do find a significant interaction, supporting the differential susceptibility hypothesis. Their estimates indicate that an extra year of education does not affect the risk of postpartum depression for mothers with 0 reactive alleles, but an extra year of education reduces the odds of postpartum depression by 18% for women with 3 or 4 alleles. Cook & Fletcher (2015) use data on 467 sibling pairs from the Wisconsin Longitudinal Study to estimate interactions between education and the APOE4 gene that has been linked to Late-onset Alzheimer's Disease. They employ sibling fixed-effects to control for confounding by

unobserved family-level factors and population stratification. They find a negative main effect of genetic risk, a positive main effect of education, and a negative GxE interaction on the probability of not experiencing a decline in cognition. The interaction estimates indicate that for high school graduates, having one copy of the APOE4 gene reduces the probability of not experiencing a cognitive decline by 25 percentage points, whereas the APOE4 gene has no effect for college graduates.

More recent studies use GRSs to measure genetic risk. Liu et al. (2015) estimate GxE interactions between (i) GRS for BMI and education and (ii) a GRS for type-2 diabetes (measured by HbA1c levels) and education, using a sample of white non-Hispanic elderly individuals from the Health & Retirement Study (HRS). As expected, they find positive main effects of genetic risk, and negative main effects of education. They find no significant interactions between education and the GRS of BMI, but do find a negative significant interaction between education and the GRS of type-2 diabetes, with stronger interactions at the higher ends of the HbA1c distribution. Liu & Guo (2015) also use the HRS, but estimate interactions between a GRS of BMI and measures of socioeconomic status (SES) throughout the life-course rather than just educational attainment. They use measures of SES in childhood (measured by father's occupation), young adulthood (measured by years of education), and middle/late adulthood (measured by household wealth). Based on these measures, they also construct a cumulative socioeconomic advantage score. They find no significant interaction between childhood SES and the GRS. However, they did find that the genetic influence on BMI is greater for individuals who experienced less socioeconomic advantage than for those who experienced more socioeconomic advantage over the life-course. Using the UK Biobank, Tyrrell et al. (2017) find a significant GxE interaction between the Townsend deprivation index (a measure of socioeconomic deprivation) and a GRS for BMI, but no significant interactions with education⁴. Amin et al. (2017) also find no significant GxE interactions between education and a GRS for BMI in the UK and Finland using the Understanding Society and Young Finns datasets.

In sum, little is known about whether education reduces genetic risk particularly with respect to BMI and obesity. Moreover, with the exception of Cook & Fletcher (2015) (who used a single gene to measure genetic risk), the existing studies do not attempt to control for sources of confounding, and therefore the estimates cannot be given a causal interpretation. This study adds to the literature by using a GRS and twins fixed-effects to control for some of the some of the sources of confounding⁵.

4. Methodology

We use data on female MZ and DZ twins to estimate equation (3) where the outcome (BMI or obesity) of twin i in pair j (Y_{ij}) is related to education (E_{ij}), a genetic risk score (GRS_{ij}), an interaction between education and the GRS ($E_{ij} * GRS_{ij}$), unobserved family-level factors (h_j), unobserved twin-specific factors (a_{ij}), and an error term (u_{ij}).

³A SNP is a DNA sequence variation occurring when a single nucleotide (A, T, C or G) in the genome differ at a single position among individuals.

⁴The Townsend deprivation index is a composite measure of deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding.

$$Y_{ij} = \gamma_1 E_{ij} + \gamma_2 GRS_{ij} + \gamma_3 (E_{ij} * GRS_{ij}) + \gamma_4 h_j + \gamma_5 a_{ij} + u_{ij} \quad (3)$$

We first treat the twins as unrelated individuals and estimate equation (3) by OLS. As noted in section 2.2 OLS estimates are likely to be biased because E_{ij} (and hence $E_{ij} * GRS_{ij}$) is correlated with a_{ij} and h_j . We then exploit the twin dimension and estimate twin fixed-effect regressions to control for the influence of a_{ij} and h_j . Twins fixed-effects exploits the idea that unobserved differences in a_{ij} and h_j are removed by differencing within twin pairs so that equation (3) can be written as:

$$\Delta Y_{ij} = \gamma_1 \Delta E_{ij} + \gamma_2 \Delta GRS_{ij} + \gamma_3 \Delta (E_{ij} * GRS_{ij}) + \gamma_4 \Delta h_j + \gamma_5 \Delta a_{ij} + \Delta u_{ij} \quad (4)$$

where Δ is the difference between twins

MZ twins are genetically identical and share the same family environment. If MZ twins are used to estimate equation (4) then the influence of shared unobserved family and twin-specific factors is eliminated ($\Delta h_j = \Delta a_{ij} = 0$). Using MZ twins we can identify the GxE interaction (γ_3), but we cannot identify the main association of the GRS (γ_2) as there is no genetic variation. DZ twins are like ordinary siblings, sharing about 50% of their genes if mating is random and the same family environment. Employing DZ twins controls for the influence of unobserved family-level factors ($\Delta h_j = 0$), and about 50% of unobserved genetic factors ($\Delta a_{ij} \neq 0$). DZ twins also control for population stratification because each twin has an equal chance of obtaining a particular gene. With DZ twins, we can estimate the main association of the GRS and the GxE interaction. We therefore focus on results using DZ twins, but we also present results using MZ twins. We acknowledge that there still may be residual confounding in twins fixed-effect estimates because they do not eliminate the influence of any unobserved factors that affect education and BMI/obesity, which are not shared by twins⁶. Twins fixed-effect estimates are therefore considered to represent quasi-causal relations (Turkheimer & Harden 2014). Nevertheless, comparisons between OLS and twins fixed-effects estimates are useful to learn the extent to which OLS estimates are biased because of a failure to account for confounding.

5. Data

The TwinsUK database consists of approximately 12,000 MZ and DZ Caucasian twins aged 16 to 100 from all over the U.K. All twins in the registry were recruited through national media campaigns and from other twin registries (Institute of Psychiatry and Aberdeen University). The majority of twins are female, because the diseases which TwinsUK was initially interested in are more common in women than men. However, the twins have not

⁵Some other studies have used sibling fixed-effects to estimate GxE interactions that control for confounding (see for example, Cook & Fletcher 2014, Pampel et al. 2015, Thompson 2014). However, these all use single genetic markers, and do not focus on education as the environment variable.

been selected for any particular disease, and they volunteer to take part in studies that cover a wide range of traits and common medical conditions (Spector et al. 2006).

As outcomes we use BMI (weight in kg/height m²), and an indicator variable for being obese (BMI ≥ 30). The height and weight measurements were taken by nurses during visits to TwinsUK. We construct a GRS for obesity based on the GWAS results of Speliotes et al. (2010). Speliotes et al. (2010) tested associations between 2.8 million SNPs and BMI for 123,865 individuals, and identified 32 SNPs that reached genome-wide significance. These SNPs explain approximately 1.45% of the variation in BMI. Of the 32 SNPs identified in Speliotes et al. (2010), only 28 were genotyped and passed quality control in TwinsUK after imputation.⁷ Genotyping was done with a combination of Illumina arrays (HumanHap3001,2, HumanHap610Q, 1M-Duo and 1.2MDuo 1M).

In a survey sent out by Bonjour et al. (2003) in 1999–2000, twins were given a list of qualifications and asked to report which qualifications they had⁸. We use responses to this question to create dummy variables indicating having (i) less than compulsory education (low vocational qualification, clerical qualification or other qualification), (ii) compulsory education (O-levels or middle vocational qualifications), (iii) some post-compulsory education (nursing, teaching, or higher vocational qualification), and (iv) a university education (university degree). The analyses in this paper are based on a sample of 288 female MZ and 1256 female DZ twins aged 25 or older when BMI measurements were collected and have non-missing information on educational attainment, BMI, and genetic risk⁹.

6. Results

Table 1 provides summary statistics. Column 1 provides statistics for the full sample of MZ and DZ twins used to estimate equation (3), while column 2 provides statistics for the sub-sample of DZ twin pairs. The average age when BMI was measured is 55 years. Average BMI is 26.14 and 18% of twins are obese. The unweighted GRS shows that on average twins have 26 risk alleles. In terms of educational attainment, 13% of twins have less than compulsory education and 35% have attained compulsory education. 31% have obtained some post-compulsory education and 20% are university graduates. Summary statistics for the sub-sample of DZ twin pairs in column 2 are similar to the those for the full sample.

Table 2 presents OLS estimates from equation (3) for the full sample of MZ and DZ twins. The weighted GRS has been standardized to have a mean of 0 and standard deviation of 1. All regressions control for age when BMI was measured. Column 1 shows that a 1 standard deviation in the GRS is associated with an increase in BMI by 0.748 kg/m². This is a relatively small association, as it implies that a 4 standard deviation increase in the GRS is

⁶This is especially the case with DZ twins because genetic differences remain, which may be correlated with education and BMI/obesity.

⁷The SNPs not genotyped were: rs2890652, rs4836133, rs4929949, and rs887912.

⁸The qualifications and associated years of schooling that are assigned by Bonjour et al. (2003) are: university, 17 years; higher vocational, 16 years; teaching, 16 years; nursing, 15 years; A-level, 14 years; middle vocational, 12 years; O-level, 12 years; low vocational, 11 years; clerical, 11 years; other qualifications, 10 years.

required to increase BMI by 11% relative to the sample mean of 26.14. The R^2 indicates that the GRS explains 8.9% of the variation in BMI. Column 2 adds indicator variables for educational attainment. The reference category is less than compulsory education. As expected, there are negative associations between education and BMI, with larger associations at the upper end of the education distribution. For example, twins with a university degree (some post-compulsory education) have on average a lower BMI of 1.702 (1.251) kg/m^2 compared to twins with less than compulsory education. Controlling for education reduces the main association of the GRS with BMI slight to 0.771. Column 3 adds the GxE interaction terms. The coefficient on the GRS now represents the marginal association of the GRS with BMI for twins with less than compulsory education. It shows that a 1 standard deviation increase in the GRS is associated with an increase in BMI of 0.812 kg/m^2 for twins who haven't completed compulsory education. While there are still significant differences in BMI by educational attainment, the marginal association of the GRS does not differ by educational attainment. All of the GxE interaction coefficients are statistically insignificant, and estimated with large standard errors. Columns 4–6 provide estimates for obesity. The pattern of results is similar to those for BMI. There are significant main association of the GRS and educational attainment with obesity. The estimates in column 5 show that a 1 standard deviation increase in the GRS increases the probability of being obese by 4.2 percentage points. This is a relatively large magnitude, as it represents a 23% increase relative to the baseline obesity rate of 18%. On average twins with a university degree are 11.8 percentage points less likely to be obese than those with less than compulsory education. This education-gradient persist in column 6. However, there are no significant differences in the marginal association of the GRS with obesity by educational attainment. All GxE interactions terms though negative are insignificant.

Table 3 presents OLS and twins fixed-effects results for BMI on the sample of 628 DZ twin pairs. comparison of estimates in columns 2 and 5 reveals that controlling for confounding through twins fixed-effects reduces the role of genetics. For example, in column 2 a 1 standard deviation increase in the GRS is associated with a higher BMI of 0.916 kg/m^2 . In comparison the corresponding twins fixed-effect estimate in column 5 is half the size, 0.487. Similarly, associations between educational attainment and BMI also decrease substantially in magnitude and become insignificant with control for confounding. The OLS estimate for university education of -2.076 in column 2 decreases by 74% to -0.553 in column 5. The twins fixed-effects estimates do not provide any evidence of any GxE interactions. In column 6, a 1 standard deviation increase in the GRS is associated with a higher BMI of 0.538 kg/m^2 for twins with less than compulsory education. This association is insignificant, as are the GxE interaction estimates.

OLS and twins fixed-effects estimates for obesity are given in table 4. OLS estimates in column 3 show that there is a 14.3 percentage point difference in the probability of being obese between twins with university education and those with less than compulsory education. There is also some indication of a GxE interaction. In particular, the coefficient on university*GRS indicates that the marginal association of the GRS with obesity is 7.3 percentage points lower for twins with a university education than for twins with less than compulsory education. This difference is marginally significant at the 10% level. Twins fixed-effects estimates in column 6 show that twins with a university education are only 2.2

percentage points less likely to be obese than twins with less than compulsory education. This difference is not statistically significant either. The marginal association of the GRS with obesity is 6.3 percentage points lower for twins with a university education, compared to twins with less than compulsory education. However, this difference is insignificant.

Table 5 presents OLS and twins fixed-effects estimates for DZ twins using only an indicator for university education. OLS estimates in column 1 show that twins with a university education have a BMI of 1.215 kg/m² lower than twins without a university education. 1 standard deviation in the GRS is associated with an increase in BMI of 0.987kg/m² for twins without a university education. The GxE interaction shows that the marginal association of the GRS for twins with a university education is not significantly different from the marginal association for twins without a university education. Twins fixed-effect estimates in column 2 show that there is no significant difference in BMI between twins with and without a university education. There is also no evidence of any GxE interactions. OLS estimates in column 3 show that there is a significant GxE interaction for obesity. Specifically, a 1 standard deviation in the GRS increases the probability of being obese by 1.5 percentage points for twins with a university education, which is 4.7 percentage points lower than the marginal association of the GRS (6.2 percentage points) for twins without a university education. Twins fixed-effects estimates in column 6 indicate that the marginal association of the GRS with obesity for twins with a university education is 1.6 percentage points. This is 3 percentage points smaller than the marginal association of the GRS with obesity for twins without a university education. However, the difference is not statistically significant, as indicated by the large standard error on the GxE interaction. Table 6 shows results where we use an indicator variable equal to 1 if the twin has post-compulsory or university education and 0 if otherwise. The pattern of results in table 6 are similar to those in table 5.

Table 7 provides twins fixed-effects estimates for 288 MZ twin pairs using different sets of indicators for educational attainment. As noted in section 4, we cannot identify the association of the GRS as MZ twins are genetically identical, but we can identify the GxE interaction terms. For BMI in column 1, while all the estimates for the education indicators are insignificant, the GxE interaction terms are significant and quite large in magnitude. For example, the interaction between university education and the GRS indicates that the BMI of twins with a university education is 2.523 kg/m² lower when the GRS is higher. However, the GxE interaction terms when using single education indicators (columns 2 and 3) are insignificant, as are the estimates for obesity in columns 4–6.

7. Conclusion

We estimate main associations of genetic risk, educational attainment, and possible GxE interactions for BMI and obesity using the TwinsUK database. Our contribution is the use of a GRS to measure genetic risk and twins fixed-effects to control for confounding. OLS estimates reveal significant main associations of genetic risk and educational attainment. The association of the GRS with BMI is small. One would require a 4 standard deviation increase in the GRS to increase BMI by about 11% of the sample mean. The association of the GRS with obesity is much larger. A 1 standard deviation increase in the GRS increases the probability of being obese by 4.1–5.2 percentage points (22%–25% relative to the

baseline obesity rate of 18%). For both BMI and obesity, the largest education gradient is observed for twins with a university education compared to twins with less than compulsory education. The OLS estimates do not provide any evidence of GxE interactions. All the GxE interaction terms are statistically insignificant. This is consistent with results in Liu et al. (2015), Amin et al. (2017) and Tyrrell et al. (2017) who also find insignificant GxE interactions. The twins fixed-effect estimates show a much smaller association of genetic risk, insignificant associations of education, and insignificant GxE interactions. Our main finding is that we do not detect any significant GxE interactions with control for confounding through twins fixed-effects. The standard errors for the GxE estimates are all larger than the point estimates, and it is difficult to draw any definite conclusions.

There are four key possible explanations for the insignificant results. First, the GRS is constructed using genes identified by meta-analyses of cohorts across time and space, and thus is a robust predictor of mean levels of BMI/obesity. However, the failure to detect GxE interactions could be because the GRS captures level effects that are common across environments rather than genetic influences that are context specific. Second, the insignificant results could be due to a relatively small sample size and lack of power. Power calculations in Conley (2015) show that when employing a GRS that explains about 2% of the variation in a phenotype, one needs less than 500 sibling pairs to detect an effect size of 0.03 for a significant GxE interaction with 60% power. We have 628 twin pairs, so our study should be adequately powered. Third, the insignificant estimates could be due to measurement error, as fixed-effects exacerbates the attenuation bias from measurement error. A fourth reason for the insignificant estimates could be a lack of within-twin pair variation in education, GRS and BMI/obesity. However, table 8 shows that there is a fair amount of variation in these measures. For the full sample of 628 DZ twin pairs in column 1, 22% of twin pairs are discordant on obesity, and the absolute within-twin pair difference in BMI is 3.77 kg/m². The within-twin pair difference in BMI is 14.5% of the sample mean, which may seem large given expectations for similarity in BMI for twins. However, both MZ and DZ twins do have relatively large differences in BMI, and these are seen in many datasets. Based on 151 DZ female twin pairs in the Midlife in United States Development Survey (MIDUS), we find that the average within-twins difference in BMI is 5.45 kg/m², which is 17% of the sample mean (26.03). MZ twins also have fairly large differences in BMI. The average within-twins difference in BMI for 296 MZ twins in the MIDUS is 2.43 kg/m², which is 9.4% of the sample mean (25.93). Amin et al. (2015) report that the within-twins difference in BMI for MZ twins in the Mid-Atlantic Twin Registry and Minnesota Twin Registry is 2.08 and 2.49 respectively. This represents 8.5% and 9.7% of the sample means respectively. In our data, the average within-twin pair difference in BMI for the 144 female MZ twin pairs is 9.7% of the sample mean. In terms of genetics, on average twin pairs have a difference of 2.62 risk alleles in the GRS. This is fairly large given that the average number of risk alleles is 25.69. Moreover, 16.72% of twin pairs have a difference of 5 or more risk alleles. In terms of educational attainment, while 44.75% of twin pairs have no difference, there are some large differences. For example, in 8.44% of twin pairs one twin has attained compulsory education while the other has completed university. The twins fixed-effects estimates are identified by twin pairs that differ in all 3 measures. We have a reasonable effective sample size to estimate the GxE regressions for BMI. There are 304 twin pairs that

are discordant on BMI, education and the GRS, and column 2 shows that there is also a reasonable amount of with-twins variation. The GxE estimates for obesity are identified by 67 twin pairs who are discordant on obesity, the GRS and education. There is also a large amount of with-twins variation for these twins (column 3).

In addition to the imprecise twins fixed-effect estimates, our study has some limitations. First, our results are unlikely to be generalizable because the sample of twins are not representative of the UK population. Second, our results are based on female twins, which is not informative about possible GxE interactions for men and any gender differences.

References

- Amin V, Behrman JR, & Kohler HP (2015). Schooling has smaller or insignificant effects on adult health in the US than suggested by cross-sectional associations: new estimates using relatively large samples of identical twins. *Social science & medicine*, 127, 181–189. [PubMed: 25110343]
- Amin V, Böckerman P, Viinikainen J, Smart MC, Bao Y, Kumari M, ... & Pehkonen J (2017). Gene-environment interactions between education and body mass: Evidence from the UK and Finland. *Social Science & Medicine*, 195, 12–16. [PubMed: 29102742]
- Becker GS, & Mulligan CB (1997). The endogenous determination of time preference. *The Quarterly Journal of Economics*, 112(3), 729–758.
- Belsky J, & Pluess M (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, 135(6), 885. [PubMed: 19883141]
- Biroli P (2015). Genetic and Economic Interaction in the Formation of Human Capital: The Case of Obesity Mimeo University of Zurich.
- Black SE, Devereux PJ, & Salvanes KG (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1), 409–439.
- Bonjour D, Cherkas LF, Haskel JE, Hawkes DD, & Spector TD (2003). Returns to education: Evidence from UK twins. *The American Economic Review*, 93(5), 1799–1812.
- Cardon LR, & Palmer LJ (2003). Population stratification and spurious allelic association. *The Lancet*, 361(9357), 598–604.
- Chandola T, Clarke P, Morris JN, & Blane D (2006). Pathways between education and health: a causal modelling approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(2), 337–359.
- Classen TJ, & Thompson O (2016). Genes and the intergenerational transmission of BMI and obesity. *Economics & Human Biology*, 23, 121–133. [PubMed: 27599025]
- Cohen AK, Rai M, Rehkopf DH, & Abrams B (2013). Educational attainment and obesity: a systematic review. *Obesity Reviews*, 14(12), 989–1005. [PubMed: 23889851]
- Conley D (2015). Genotyping a new, national household panel study: White paper prepared for NSF-sponsored Conference, May 2014. *Journal of Economic and Social Measurement*, 40(1–4), 375–395.
- Conley D, & Rauscher E (2013). Genetic interactions with prenatal social environment: Effects on academic and behavioral outcomes. *Journal of Health and Social Behavior*, 54(1), 109–127. [PubMed: 23475742]
- Cook CJ, & Fletcher JM (2014). Interactive effects of in utero nutrition and genetic inheritance on cognition: New evidence using sibling comparisons. *Economics & Human Biology*, 13, 144–154. [PubMed: 24172871]
- Cook CJ, & Fletcher JM (2015). Can education rescue genetic liability for cognitive decline? *Social Science & Medicine*, 127, 159–170. [PubMed: 25074513]
- Datar A, & Sturm R (2006). Childhood overweight and elementary school outcomes. *International journal of obesity*, 30(9), 1449. [PubMed: 16534518]

- Ellis BJ, Boyce WT, Belsky J, Bakermans-Kranenburg MJ, & Van IJzendoorn MH (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. *Development and Psychopathology*, 23(1), 7–28. [PubMed: 21262036]
- Figlio D, Guryan J, Karbownik K, & Roth J (2014). The effects of poor neonatal health on children’s cognitive development. *American Economic Review*, 104(12), 3921–55. [PubMed: 29533575]
- Fletcher JM (2010). Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects. *Health Economics*, 19(4), 466–484. [PubMed: 19382102]
- Grossman M (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2), 223–255.
- Iliadou A, Cnattingius S, & Lichtenstein P (2004). Low birthweight and type 2 diabetes: a study on 11 162 Swedish twins. *International journal of epidemiology*, 33(5), 948–953. [PubMed: 15166205]
- Kaestner R, & Grossman M (2009). Effects of weight on children’s educational achievement. *Economics of Education Review*, 28(6), 651–661.
- Kenkel DS (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(2), 287–305.
- 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. [PubMed: 27231400]
- Liu SY, Walter S, Marden J, Rehkopf DH, Kubzansky LD, Nguyen T, & Glymour MM (2015). Genetic vulnerability to diabetes and obesity: Does education offset the risk?. *Social Science & Medicine*, 127, 150–158, 80–94. [PubMed: 25245452]
- Lundborg P, & Stenberg A (2010). Nature, nurture and socioeconomic policy—What can we learn from molecular genetics?. *Economics & Human Biology*, 8(3), 320–330. [PubMed: 20833117]
- Maes HH, Neale MC, & Eaves LJ (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics*, 27(4), 325–351. [PubMed: 9519560]
- Mitchell C, Notterman D, Brooks-Gunn J, Hobcraft J, Garfinkel I, Jaeger K, ... & McLanahan S (2011). Role of mother’s genes and environment in postpartum depression. *Proceedings of the National Academy of Sciences*, 108(20), 8189–8193.
- Pampel FC, Boardman JD, Daw J, Stallings MC, Smolen A, Haberstick BC, ... & Conger RD (2015). Life events, genetic susceptibility, and smoking among adolescents. *Social Science Research*, 54, 221–232. [PubMed: 26463545]
- Plomin R, Haworth CM, & Davis OS (2009). Common disorders are quantitative traits. *Nature Reviews Genetics*, 10(12), 872–878.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, & Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), 904–909. [PubMed: 16862161]
- Rose G, & Marmot MG (1981). Social class and coronary heart disease. *Heart*, 45(1), 13–19.
- Royer H (2009). Separated at girth: US twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1), 49–85.
- Sabia JJ (2007). The effect of body weight on adolescent academic performance. *Southern Economic Journal*, 871–900.
- Spector TD, & Williams FM (2006). The UK adult twin registry (TwinsUK). *Twin Research and Human Genetics*, 9(6), 899–906. [PubMed: 17254428]
- Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, ... & Randall JC (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics*, 42(11), 937–948. [PubMed: 20935630]
- Thompson O (2014). Economic background and educational attainment the role of gene-environment interactions. *Journal of Human Resources*, 49(2), 263–294.
- Turkheimer E, Harden KP. Handbook of research methods in social and personality psychology. In: Resi HT, Judd CM, editors. *Behavior genetic research methods: Testing quasi-causal hypotheses using multivariate twin data* 2nd. Cambridge, UK: Cambridge University Press: 2014 p.159–189

- Tyrrell J, Wood AR, Ames RM, Yaghootkar H, Beaumont RN, Jones SE, ... & Joost S (2017). Gene–obesogenic environment interactions in the UK Biobank study. *International Journal of Epidemiology*, 46(2), 559–575. [PubMed: 28073954]
- Visscher PM, Hill WG, & Wray NR (2008). Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics*, 9(4), 255–266.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlight

- We investigate GxE interactions between education and genetic risk for obesity.
- We use twins fixed-effects to control for confounding.
- GxE interactions are insignificant but imprecisely estimated.

Table 1:

Summary Statistics

	MZ + DZ Twins	DZ Twin Pairs
	(1)	(2)
Age	55.51 (12.49)	55.12 (12.25)
Female	1.00 (0.00)	1.00 (0.00)
MZ	0.17 (0.39)	0.00 (0.00)
BMI	26.14 (4.92)	26.12 (4.81)
Obese	0.18 (0.38)	0.18 (0.38)
Less than compulsory education	0.13 (0.34)	0.13 (0.33)
Compulsory education	0.35 (0.48)	0.34 (0.48)
Some post-compulsory education	0.31 (0.46)	0.32 (0.46)
University education	0.21 (0.40)	0.21 (0.40)
GRS (unweighted)	25.67 (3.22)	25.69 (3.19)
GRS (weighted)	3.76 (0.52)	3.76 (0.52)
N	1544	1256

Notes: Standard deviations in parentheses.

Table 2:

OLS estimates of main associations of education, genetic risk, and GxE interactions using MZ and DZ Twins

	BMI	BMI	BMI	Obese	Obese	Obese
	(1)	(2)	(3)	(4)	(5)	(6)
Genetics						
GRS	0.748 ***	0.771 ***	0.812 **	0.041 ***	0.042 **	0.063 **
	(0.145)	(0.144)	(0.392)	(0.010)	(0.103)	(0.029)
Educational Attainment						
Compulsory education		-0.494	-0.487		-0.051	-0.05
		(0.457)	(0.459)		(0.037)	(0.038)
Post-compulsory education		-1.251 ***	-1.249 **		-0.084 **	-0.084 **
		(0.457)	(0.459)		(0.037)	(0.038)
University education		-1.702 ***	-1.696 ***		-0.118 ***	-0.118 ***
		(0.498)	(0.502)		(0.039)	(0.039)
GxE						
Compulsory education *GRS			0.081			-0.004
			(0.434)			(0.003)
Post-compulsory education *GRS			-0.021			-0.025
			(0.421)			(0.032)
University education *GRS			-0.318			-0.057
			(0.463)			(0.034)
R-squared	0.089	0.102	0.103	0.062	0.071	0.073
N	1544	1544	1544	1544	1544	1544

Notes: All regressions control for age when BMI was measured. The omitted education reference category is less than compulsory education. The weighted GRS has been standardized to have a mean of 0 and standard deviation of 1. The within R^2 from the twin-differenced regression (equation 4) is reported in columns 4–6. Robust standard errors clustered by twin pair are reported in parentheses

* statistically significant at the 0.10 level

** at the 0.05 level

*** at the 0.01 level.

Table 3:

OLS and Twins Fixed-Effect estimates of main associations of education, genetic risk, and GxE interactions for BMI using DZ Twins

	BMI	BMI	BMI	BMI	BMI	BMI
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	FE	FE	FE
Genetics						
GRS	0.881 *** *	0.916 ***	1.117 **	0.475 **	0.487 **	0.538
	(0.152)	(0.152)	(0.427)	(0.188)	(0.189)	(0.537)
Educational Attainment						
Compulsory education		-0.600	-0.603		-0.141	-0.139
		(0.491)	(0.494)		(0.593)	(0.593)
Post-compulsory education		-1.343 **	-1.356 **		-0.200	-0.230
		(0.481)	(0.482)		(0.612)	(0.611)
University education		- 2.076 ***	- 2.063 ***		-0.553	-0.572
		(0.519)	(0.520)		(0.712)	(0.710)
GxE						
Compulsory education *GRS			-0.125			-0.09
			(0.475)			(0.575)
Post-compulsory education *GRS			-0.123			0.125
			(0.470)			(0.603)
University education *GRS			-0.591			-0.338
			-0.494			-0.716
R-squared	0.098	0.115	0.117	0.116	0.118	0.120
N	1256	1256	1256	1256	1256	1256

Notes: All regressions control for age when BMI was measured. The omitted education reference category is less than compulsory education. The weighted GRS has been standardized to have a mean of 0 and standard deviation of 1. The within R^2 from the twin-differenced regression (equation 4) is reported in columns 4–6. Robust standard errors clustered by twin pair are reported in parentheses

* statistically significant at the 0.10 level

** at the 0.05 level

*** at the 0.01 level.

Table 4:

OLS and Twins Fixed-Effect estimates of main associations of education, genetic risk, and GxE interactions for Obesity using DZ twins

	Obese	Obese	Obese	Obese	Obese	Obese
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	FE	FE	FE
Genetics						
GRS	0.052 ***	0.054 ***	0.088 **	0.040 **	0.040 **	0.079
	(0.011)	(0.011)	(0.033)	(0.018)	(0.018)	(0.050)
Educational Attainment						
Compulsory education		-0.068 *	-0.068		-0.02	-0.021
		(0.041)	(0.042)		(0.058)	(0.059)
Post-compulsory education		-	-		-	-
		0.095 **	-0.095 **		-0.008	-0.009
		(0.042)	(0.042)		(0.063)	(0.062)
University education		-	-		-	-
		0.144 ***	0.143 ***		-0.022	-0.022
		(0.043)	(0.043)		(0.067)	(0.067)
GxE						
Compulsory education *GRS			-0.019			-0.049
			(0.038)			(0.054)
Post-compulsory education *GRS			-0.039			-0.031
			(0.037)			(0.053)
University education *GRS			-0.073 *			-0.063
			(0.038)			(0.057)
R-squared		0.068	0.079	0.082	0.042 0.043	0.046
N		1256	1256	1256	1256	1256

Notes: All regressions control for age when BMI was measured. The omitted education reference category is less than compulsory education. The weighted GRS has been standardized to have a mean of 0 and standard deviation of 1. The within R^2 from the twin-differenced regression (equation 4) is reported in columns 4–6. Robust standard errors clustered by twin pair are reported in parentheses

* statistically significant at the 0.10 level

** at the 0.05 level

*** at the 0.01 level.

Table 5:

OLS and Twins Fixed-Effect estimates of main associations of education, genetic risk, and GxE interactions using DZ twins and measuring education using an indicator for university education

	BMI	BMI	Obese	Obese
	(1)	(2)	(3)	(4)
	OLS	FE	OLS	FE
Genetics				
GRS	0.987 ***	0.556 **	0.062 ***	0.046 **
	(0.175)	(0.218)	(0.013)	(0.021)
Educational Attainment				
University education	-1.215 **	-0.392	-0.072 *** - 0.011	
	(0.313)	(0.380)	(0.024)	(0.037)
GxE				
University education *GRS	-0.461	-0.363	-0.047 **	-0.03
	(0.304)	(0.425)	(0.021)	(0.034)
R-squared	0.109	0.118	0.076	0.043
N	1256	1256	1256	1256

Notes: All regressions control for age when BMI was measured. The omitted education reference category is no university education. The weighted GRS has been standardized to have a mean of 0 and standard deviation of 1. The within R^2 from the twin-differenced regression (equation 4) is reported in columns 1–4. Robust standard errors clustered by twin pair are reported in parentheses

* statistically significant at the 0.10 level

** at the 0.05 level

*** at the 0.01 level.

Table 6:

OLS and Twins Fixed-Effect estimates of main associations of education, genetic risk, and GxE interactions using DZ twins and measuring education using an indicator for post-compulsory education or higher

	BMI	BMI	Obese	Obese
	(1)	(2)	(3)	(4)
	OLS	FE	OLS	FE
Genetics				
GRS	1.025 ***	0.45	0.074 ***	0.041
	(0.216)	(0.281)	(0.016)	(0.027)
Educational Attainment				
Post-compulsory education	-1.143 **	-0.185	-0.063 ***	
or higher	(0.284)	(0.361)	(0.024)	(0.034)
GxE				
Post-compulsory education *GRS	-0.214	-0.061	-0.039 *	-0.004
	(0.253)	(0.345)	(0.020)	(0.029)
R-squared	0.111	0.117	0.077	0.042
N	1256	1256	1256	1256

Notes: All regressions control for age when BMI was measured. The omitted education reference category is less than post-compulsory education.

The weighted GRS has been standardized to have a mean of 0 and standard deviation of 1. The within R^2 from the twin-differenced regression (equation 4) is reported in columns 1–4. Robust standard errors clustered by twin pair are reported in parentheses

* statistically significant at the 0.10 level

** at the 0.05 level

*** at the 0.01 level.

Table 7:

Twins Fixed-Effect estimates for BMI and Obesity using MZ twins

	BMI	BMI	BMI	Obese	Obese	Obese
	(1)	(2)	(3)	(4)	(5)	(6)
	FE	FE	FE	FE	FE	FE
Educational Attainment						
Compulsory education	-0.527			- 0.147 *		
	(0.698)			(0.081)		
Post-compulsory education	-0.653			-0.035		
	(0.714)			(0.079)		
Post-compulsory education		-0.191			0.094 *	
or higher		(0.515)			(0.053)	
University education	-0.607		-0.051	-0.017		0.043
	(0.892)		(0.624)	(0.099)		(0.071)
GxE						
Compulsory education *GRS	-0.193 **			-0.062		
	(0.712)			(0.086)		
Post-compulsory education *GRS	-2.094 ***			-0.052		
	(0.629)			(0.073)		
Post-compulsory education or		-0.609			-0.013	
higher *GRS		(0.426)			(0.028)	
(.028)						
University education *GRS	- 2.523 ***		-0.48	-0.128		-0.076
	(0.846)		(0.562)	(0.101)		(0.072)
R-squared	0.07	0.016	0.009	0.046	0.015	0.009
N	288	288	288	288	288	288

Notes: All regressions control for age when BMI was measured. The weighted GRS has been standardized to have a mean of 0 and standard deviation of 1. The within R^2 from the twin-differenced regression (equation 4) is reported in columns 1–6. Robust standard errors clustered by twin pair are reported in parentheses

* statistically significant at the 0.10 level

** at the 0.05 level

*** at the 0.01 level.

Table 8:

Within-Twin Differences for DZ twins

	All DZ Twins	DZ twins discordant on education, BMI, & GRS	DZ twins discordant on education, obesity, & GRS
	(1)	(2)	(3)
Absolute within-twin pair difference in BMI	3.77 (3.38)	3.68 (3.34)	7.34 (4.68)
Absolute within-twin pair difference in obesity	0.22 (0.42)	0.22 (0.42)	1.00 (0.00)
Absolute within-twin pair difference in GRS	2.62 (1.94)	2.98 (1.76)	3.10 (1.92)
Percentage of twin pairs with			
0 difference in GRS	12.1	0	0
1 risk allele difference in GRS	21.66	23.36	25.37
2 risk allele difference in GRS	19.9	23.03	23.88
3 risk allele difference in GRS	17.52	22.04	13.43
4 risk allele difference in GRS	12.1	12.5	13.43
5 or more risk allele difference in GRS	16.72	19.08	17.91
% of twin pairs with			
no difference in education	44.75	0	0
less than compulsory vs compulsory	7.48	14.14	14.93
less than compulsory vs post-compulsory	6.37	12.5	22.39
less than compulsory vs university	0.96	1.64	2.99
compulsory vs post-compulsory	17.68	29.93	29.85
compulsory vs university	8.44	15.46	10.45
post-compulsory vs university	14.33	26.32	19.4
Number of pairs	628	304	67

Notes: Standard deviations in parentheses.