Author(s): Pehkonen, Jaakko; Viinikainen, Jutta; Böckerman, Petri; Lehtimäki, Terho; Pitkänen, Niina; Raitakari, Olli

Title: Genetic endowments, parental resources and adult health : Evidence from the Young Finns Study

Year: 2017

Version:

Please cite the original version:

Pehkonen, J., Viinikainen, J., Böckerman, P., Lehtimäki, T., Pitkänen, N., & Raitakari, O. (2017). Genetic endowments, parental resources and adult health : Evidence from the Young Finns Study. Social Science and Medicine, 188, 191-200. https://doi.org/10.1016/j.socscimed.2017.04.030

Accepted Manuscript

Genetic endowments, parental resources and adult health: Evidence from the Young Finns Study

Jaakko Pehkonen, Jutta Viinikainen, Petri Böckerman, Terho Lehtimäki, Niina Pitkänen, Olli Raitakari

PII: S0277-9536(17)30260-5

DOI: 10.1016/j.socscimed.2017.04.030

Reference: SSM 11187

To appear in: *Social Science & Medicine*

Received Date: 9 January 2017

Revised Date: 18 April 2017

Accepted Date: 20 April 2017

Please cite this article as: Pehkonen, J., Viinikainen, J., Böckerman, P., Lehtimäki, T., Pitkänen, N., Raitakari, O., Genetic endowments, parental resources and adult health: Evidence from the Young Finns Study, *Social Science & Medicine* (2017), doi: 10.1016/j.socscimed.2017.04.030.

Genetic endowments, parental resources and adult health: Evidence from the Young Finns Study

Jaakko Pehkonen^*, Jutta Viinikainen*, Petri Böckerman**, Terho Lehtimäki***, Niina Pitkänen***, Olli Raitakari****

* School of Business and Economics, University of Jyvaskyla, Finland.

**Turku School of Economics, Labour Institute for Economic Research, Helsinki, Finland and IZA.

*** Department of Clinical Chemistry, Fimlab Laboratories and School of Medicine, University of Tampere, Finland.

**** Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku and Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland.

^Corresponding author: Phone: +358-50-3732133. Fax: +358-14-2535 7332. E-mail: jaakko.k.pehkonen@jyu.fi

Genetic endowments, parental resources and adult health: Evidence from the Young Finns Study

Abstract

This paper uses longitudinal survey data linked to administrative registers to examine socioeconomic gradients in health, particularly whether the effects of genetic endowments interact with the socioeconomic resources of the parental household. We find that genetic risk scores contribute to adult health measured by biomarkers. This result is consistent with the findings from genome-wide association studies. Socioeconomic gradients in health differ based on biomarker and resource measures. Family education is negatively related to obesity and the waist-hip ratio, and family income is negatively related to low-density lipoprotein cholesterol and triglyceride levels. Parental resources do not modify the effects of genetic endowment on adult health. However, there is evidence for gene-family income interactions for triglyceride levels, particularly among women.

Keywords: genetic risk scores, biomarkers, adult health, parental resources, genome-wide association studies

1. Introduction

Health outcomes are multifactorial and reflect genetic and environmental influences (Belsky et al., 2013; Hunter, 2005). According to the epidemiological literature, the risks conveyed by specific genotypes may also depend on the exposure levels to environmental factors (Caspi, 2002, 2003; Mattei et al., 2012). Recent research on adult obesity serves as an example of this multidimensionality. For instance, physically active individuals who carry the obesity-promoting gene may have a 30% lower risk of obesity than physically inactive individuals who carry the same gene (Kilpeläinen et al., 2011).

An intriguing but under-researched aspect of gene-environment interactions is the effect of childhood socioeconomic status (SES) on health. This topic is important for three related reasons. First, research in economics has documented significant associations between parental resources and later life outcomes of children (Björklund and Salvanes, 2010; Björklund and Jäntti, 2012; Black and Devereaux, 2010). Second, evidence regarding whether parental resources modify (i.e., mitigate or exacerbate) health risks related to genetic endowments is scant and focused on the role of "the home environment" (Bouzigon et al., 2007; Foley et al., 2004). Third, research on gene-parental correlations (rGE) and gene-parental interactions (G*E) provide useful insights into the scope of policy interventions (Conley, 2016; Thompson, 2014).

We use longitudinal survey data (Young Finns Study; YFS) that are linked to comprehensive administrative registers to examine the relationships between several health traits, parental resources and genetic variants identified in genome-wide association studies (GWAS). We focus on five health outcomes: body mass index (BMI),

waist-to-hip ratio (WHR), triglyceride level, high-density lipoprotein (HDL) cholesterol level and low-density lipoprotein (LDL) cholesterol level. These traits, particularly abdominal (central) obesity, low HDL cholesterol levels and high triglyceride levels, are factors that elevate the risk of heart disease (Boronat et al., 2009; Mendis, 2011). To provide evidence for gene-environment interactions, we examine whether the effects of genetic endowments interact with the socioeconomic resources of the parental household. As shown in twin and adoption studies, children's outcomes driven by genetic endowments may be influenced by paternal investments (Björklund et al., 2006; Turkheimer et al., 2003). The modifying role has also been detected in studies that use genetic data, of which Caspi et al. (2002, 2003) was the first to study depression and antisocial behaviour; see Lundborg and Stenberg (2010) for a comprehensive review.

This paper documents three findings. First, both parental resources and genetic risk scores contribute to health outcomes. This result confirms the findings of prior GWASs (Heid et al., 2013; Hernesniemi et al., 2015; Speliotes et al., 2010; Teslovich et al., 2011) and the earlier patterns of SES gradients in health in Finland (Valkonen et al., 1997; Tarkiainen et al., 2012) using the YFS data. Second, SES gradients in health differ according to biomarker and SES measure. Family education is negatively related to obesity (BMI and WHR), and family income is negatively related to LDL cholesterol and triglyceride levels. Third, although evidence indicates that gene-family income interactions can influence triglyceride levels, particularly in the older female cohorts of the YFS, parental resources do not modify genetic effects.

2. Context, data and descriptive statistics

2.1 Context

As in other developed countries, notable differences in health are observed in Finland according to SES level. Despite universal access to health care in Finland, a better education level is strongly related to improved health status. Therefore, the level of education is positively related to both life expectancy and disability-free life expectancy (Valkonen et al., 1997). For example, 2007 data (Tarkiainen et al., 2012) reveal a disparity in life expectancy at age 35 between the highest and lowest income quintiles of 12.5 years for men and 6.8 years for women. The corresponding differences between the highest and lowest occupational social classes (managers versus workers) were 6.1 years among men and 3.5 years among women. The differences also increased in Finland over the period 1988-2007. More highly educated individuals are also less likely to be heavy drinkers of alcohol or smokers than poorly educated individuals (Böckerman and Maczulskij, 2016). Overall, the disparities in health by SES are not substantially smaller in the Nordic countries, such as Finland, than in other countries in Europe (Mackenbach et al., 1997; Kunst et al., 2005).

2.2 Linked YFS data

This study links data from three sources: the Cardiovascular YFS; the Finnish Longitudinal Employer-Employee Data (FLEED) of Statistics Finland (SF); and the Longitudinal Population Census (LPC) of SF. The linkage of FLEED and LPC to the YFS data is possible using personal identifiers. The identifiers are exactly matched; i.e., there are no misreported ID codes.

The YFS that began in 1980 is the largest running follow-up study in Europe to evaluate cardiovascular risk factors from childhood to adulthood. The YFS collected data through questionnaires, physical measurements, and blood tests. Subjects in six age cohorts (aged 3, 6, 9, 12, 15 and 18 years) were randomly chosen from the five university hospital districts of Finland using the national population register. Eight waves of data were collected starting with the baseline in 1980, with response rates ranging between 60 and 80%. The YFS is not nationally representative with respect to Finland's total population but is representative for the selected six age cohorts (Raitakari et al., 2008). According to the Income Distribution Survey of SF, the average earnings level in Finland was 28,598 euros in 2010, and the Gini coefficient was 26.6, and the corresponding values from the YFS are 27,383 euros and 37.9, respectively. The YFS provides information on biomarkers measured at several points in time, physical activity and diet at the age of 15, birth weight, genetic risk scores, and region of birth. The second data set, i.e., the LPC, served as the source of information on parental education and income as well as father's field of occupation in 1980. The third data set, i.e., FLEED, was used to measure participant education, income and native language. The LPC and FLEED data originate in administrative registers maintained by the SF.

The analysis focuses on five biomarkers: body mass index, waist-to-hip ratio, triglycerides, HDL cholesterol, and LDL cholesterol. The measures are based on professional examinations conducted at health centres in 2001, 2007 and 2011; see Raitakari et al. (2003) for measurement details. In 2001 (2011), the youngest participant was 24 (34) years of age, and the oldest participant 39 (49) years. Thus, the average age in the sample is 36 years. The correlations between the cross-sections as well as between the biomarkers are strong. For example, between the cross-sections of

2001 and 2011, the correlations vary from 0.849 (for BMI) to 0.548 (for triglycerides). The correlations between biomarkers are strong for each pair ($p<0.01$), being strongest for BMI and WHR ($r=0.578$) and weakest for HDL and LDL cholesterols ($r=-0.089$). To alleviate possible measurement errors and short-term variations in health, we use averages of three measurements (Gatto, 2004). This approach also maximizes the sample size because the average is calculated based on two or even one observation in the case of missing values.

Genetic risk scores (GRSs) are constructed using risk variants identified in GWASs. The use of genetic scores instead of single nucleotide polymorphisms (SNPs) has two advantages. First, it provides parsimonious representation of the data and thus increases statistical power of estimation. Second, it reduces the risk that any individual SNP will bias the estimates via an alternative biological pathway (Palmer et al., 2012). We use the unweighted genetic risk score, which was calculated as the sum of the genotyped risk alleles or imputed allele dosages carried by an individual; see Böckerman et al. (2017), who document the benefits of the unweighted risk scores as opposed to the weighted scores in the YFS data.

The GRS for BMI is based on 32 SNPs associated with BMI at the genome-wide level, according to Speliotes et al. (2010). The GRS for WHR is based on 16 SNPs (Heid et al., 2010), the GRS for triglyceride on 41 SNPs (Hernesniemi et al., 2015), the GRS for HDL cholesterol on 38 SNPs (Teslovich et al., 2011), and the GRS for LDL cholesterol risk score on 58 SNPs (Hernesniemi et al., 2015). The correlations between risk scores and biomarkers in the YFS data are consistent with the GWAS evidence: they are statistically significant ($p<0.01$) for all biomarkers. The relationship is strongest for LDL cholesterol

(r=0.233) and weakest for WHR (r=0.055). For BMI r=0.133, for triglycerides r=0.157, and for HDL r=0.166.

Two measures for family resources are used: parents' university-level education (SES_1) and the log of family income (SES_2). Income is measured as the logarithm of family earnings in 1980, and the indicator for family education equals one if at least one of the parents has a university education. The genetic risk scores and the measures of SES are not correlated ($p>0.2$ in all cases).

Potential omitted variable bias is alleviated by using measures that potentially correlate with health outcomes and parental resources. First, we use predetermined background covariates: participant birth month and birth year, gender, region of birth, and native language. Second, we augment the regression analysis with covariates that capture the participants' initial health endowment, health-related behaviour, and investments in human capital (Conti and Heckman, 2010). Early health endowment is measured using birth weight. The YFS data on birth weight are well representative of the Finnish population (NIHW, 2014). The median weight in the YFS data is 3.50 kg (SE = 0.545), and the fraction of low birth weight children (less than 2.5 kg) is 3.5%. For human capital, we use an indicator based on the highest obtained degree in 2010 and participants' own income in 2001. The share of participants with tertiary-level education (33.0%) also matches well with the population. Health-related behaviours are measured using information on participants' physical activity and diet at the age of 15 (Mansikkaniemi et al., 2012). The measures are based on physical activity (seven-point scale: 1=never to 7=every day), its intensity (three-point scale: 1=no sweating to 3=extensive sweating), and the frequency of fruit and vegetable consumption, rated on a four-point scale (1=less than once a week to 4=every day).

2.3 Descriptive statistics on gene-health-SES gradients

Table 1 provides descriptive statistics, which indicate whether participants' health, genetic risks, and background characteristics are equally distributed across parental SES. Three findings are notable. First, health-SES gradients are significant. The gradients are notable for family income ($p<0.01$ for BMI, WHR and HDL; $p<0.05$ for LDL; $p<0.10$ for triglycerides). For family education, the gradients are found for BMI and WHR ($p<0.01$). Second, the participant characteristics exhibit strong SES gradients ($p<0.01$), excluding gender and birth weight for family education. Third, the sample means for the genetic risk scores do not differ according to SES ($p>0.05$ for all GRS). Additionally, the risk score distributions are similar for "high" or "low" parental SES; see Appendix 1, which compares the participants by family income and family education: top 20% versus bottom 20% (Figure A1) and high versus low education (Figure A2).

3. Empirical methods

3.1 OLS models

We explain adult health outcomes by regressing health outcomes (biomarkers, j=1…5) on parental income (SES_1), parental education (SES_2) and on a corresponding risk score (GRS_j). To alleviate omitted variable bias, we augment the regressions with predetermined covariates (vector **X**) and covariates that capture observed heterogeneity among the participants (vector **Z**). The OLS models are of the following form (omitting subscripts for individuals):

$$(1)\ Health_j = \alpha_{0j} + \alpha_{1j}(GRS)_j + \alpha_{2j}(SES_1)_j + \alpha_{3j}(SES_2)_j + \mathbf{X}_j{*}\boldsymbol{\theta}_1 + \mathbf{Z}_j{*}\boldsymbol{\theta}_2 + \varepsilon_j$$

The vector X consists of predetermined measures for birth month, birth year, gender, region of birth, and native language. The vector **Z** consists of measures for human capital inputs, including a measure for initial health endowment (birth weight), health-related behaviours at the age of 15 (diet, physical activity), human investments in education, and the income of the participants in 2001. To investigate whether the effect of genotype varies by SES, we estimate an interaction model for each health outcome (j=1…5) and both SES measures (k=1,2). The OLS models have the following structure:

$$(2)\ \text{Health}_j = \beta_{1j} + \beta_{1j}(\text{GRS})_j + \beta_{2jk}(\text{SES})_{jk} + \beta_{3jk}(\text{SES})_{jk*}(\text{GRS})_j + \mathbf{X}_j*\boldsymbol{\psi}_1 + \mathbf{Z}_j*\boldsymbol{\psi}_2 + \mu_j$$

where, as above, **X** and **Z** describe the vectors of the covariates and an estimate for β_{3jk} captures the effect of how parental resources modify the gene-health gradient.

3.2 Model extensions and robustness

We check the robustness of the results in four ways. First, we estimate the models for BMI, WHR and HDL cholesterol using a weighted genetic risk score. For LDL cholesterol and triglycerides, information on weighted risk scores is not available in the YFS. A weighted genetic risk score is calculated as the sum of imputed allele dosages carried by an individual each multiplied by the effect size (Speliotes et al., 2010, for BMI; Heid et al., 2010, for WHR; and Teslovich et al., 2011, for HDL cholesterol). Second, we estimate the models using yearly observations (2001, 2007, and 2011) for biomarkers instead of averages based on three measurements. Third, we estimate logit models by grouping dependent variables. The thresholds are as follows: Overweight: BMI≥25; Obesity: BMI≥30; Central obesity: WHR>0.85 for females, WHR>0.9 for males; high serum triglycerides: triglyceride≥1.7 mmol/l; low HLD cholesterol: HDL≤0.9 mmol/l for males,

HDL≤1.0 mmol/l for females; elevated LDL cholesterol: LDL>3.36 mmol/l. The thresholds for obesity, central obesity (WHR), high serum triglycerides and low HDL cholesterol levels are based on the thresholds of the World Health Organization (WHO), which can be used to diagnose metabolic syndrome (Alberti and Zimmet, 1998). The definitions of overweight and high LDL cholesterol level are based on WHO classifications (WHO, 2017a; WHO, 2017b). Fourth, we employ IV techniques to account for the potential endogeneity of parental resources. To this end, we use the average wage rate of the father's industry of occupation as an instrument for family income and augment the OLS models (1) and (2) using first-stage regressions, where we assume that conditions for valid instruments prevail, i.e., Cov (SES_i, Z_i) ≠ 0 and Cov (Z_i, ε_j)=0.

4. Empirical results of the determinants of health outcomes

Tables 2-4 report the estimates of the regressions that explain health outcomes. Table 2 summarizes the results of the relationship for polygenetic risk score, family resources and health. Table 3 summarizes the estimates for the gene-SES interaction terms. Table 4 provides the results of sensitivity checks, including the estimates for weighted risk scores, cross-sections and the logit and IV models.

4.1 Genetic risk scores and health outcomes

Table 2 documents three specifications for each health outcome: i) the baseline model, which only includes the predetermined covariates; ii) the augmented baseline model, which includes parental resources (SES1 and SES2); and iii) the full model, which includes all additional covariates. Because of the construction of the risk score variables,

the expected regression coefficients are positive in all cases. That is, higher GRS is related to higher levels of BMI, WHR, HDL and LDL cholesterol and triglycerides and lower levels of HDL cholesterol.

The OLS results reveal two key findings. First, the point estimate for the genetic risk score is significant for all biomarkers in all specifications ($p<0.05$), and the strength of the relationship is robust to the inclusion of additional covariates. For example, the increase in the coefficient of the risk score in the LDL regression from 0.727 (SE=0.064) to 0.791 (SE=0.071) stems solely from the variation in the sample size: the re-estimation of specification 1 with N=1558 yields approximately the same estimate (0.787) with a standard error of 0.071. For triglycerides, the corresponding estimates (i.e., full sample with no covariates, smaller sample with all covariates, and smaller sample with no covariates) were 0.905 (SE=0.096), 0.852 (SE=0.114), and 0.837 (SE=0.113), respectively.

Second, genetic risk scores, i.e., candidate genes, only account for a small fraction of the susceptibility to health traits. The magnitude of the association also varies across biomarkers. For BMI, the increase in the explanatory power associated with the risk score variable is 1.8 percentage points (%-points) as measured by an incremental R^2. The effect of the risk score is strongest for LDL cholesterol (5.3%-points) and lowest for WHR (0.20%-points). For BMI, a one standard deviation increase in the risk score is associated with an increase of approximately 2.3% in the dependent variable. For WHR, the corresponding impact is approximately 0.5%, for LDL cholesterol 5.6%, for HDL cholesterol 3.9%, and for triglycerides 9.3%.

4.2 Family resources and health outcomes

Parental resources contribute to health, and these associations vary by resource measure and biomarker (Table 2). Family resources are jointly significant (F-statistics $p<0.05$) and robust in all but two cases: HDL cholesterol ($p<0.35$) and triglycerides ($p<0.22$). For BMI and WHR, the results reveal a significant negative association with family education but not with family income. We find a robust negative association between LDL cholesterol and family income, with a point estimate of approximately -0.02 ($p<0.01$). We also document a relationship between triglycerides and family income (0.025; $p<0.05$). However, the inclusion of the covariates makes this association insignificant ($p>0.10$), although this result may reflect the smaller subsample.

4.3 Role of covariates

The covariates for participants' initial health endowment, health-related behaviours in adolescence and human capital do not attenuate the gene-health gradients: the relationships are robust across specifications (see specifications 3, 6, 9, 12 and 15 in Table 2). Participants' own income and intensity of physical activity are associated with HDL cholesterol level ($p<0.01$); diet is associated with LDL cholesterol level ($p<0.01$); birth weight and intensity of physical activity are associated with triglycerides ($p<0.01$); participants' own education, birth weight and diet are associated with obesity ($p<0.01$); and participants' own education and intensity of physical activity ($p<0.05$) are associated with WHR. Overall, the findings are consistent with prior studies that have documented associations between health outcomes, education and health-related behaviours (Conti and Heckman, 2010; Lechner, 2009).

4.4 Gene-parental resource interactions

Table 3 examines whether the effect of genetic risk varies according to SES. The results for the full sample (column 1 for income and column 6 for education) show that parental resources play a limited role in modifying the effects of genetic endowment. Overall, we find no evidence of statistically significant or economically relevant interactions between parental resources and genetic risk scores. The only specification that displays a significant association is the model based on triglycerides when parental resources are measured using family income (column 1). The estimate (-0.44; $p<0.01$) indicates that high family income attenuates disadvantaged genetic background. Higher family income may also modify the effects of genetic background in the case of HDL cholesterol: the interaction coefficient was positive but significant only at the 10% level.

Health disparities may be gender- and cohort-dependent (Andersson, 2016). To investigate this possibility, we estimate gene-SES gradients for the gender and age-based participant cohorts. These results are documented in columns 2-5 and 7-10. The results are consistent with the aggregate analysis. There is evidence that high family income modifies poor genetic background for triglycerides, particularly for females ($p<0.05$). However, overall, we do not detect significant interaction effects. In the Supplementary Web Appendix, we provide further robustness analyses for the interaction effects (results based on weighted risk score and cross-sections, logit, and IV results; Tables SA1-SA10). Apart from triglycerides, the modifying role of parental background appears to be weak.

4.5 Robustness

Table 4 reports the results of a sensitivity analysis for one biomarker (BMI). Estimates for other biomarkers are reported in the Supplementary Web Appendix (Tables SA11-SA14). Columns 1 and 2 show that the estimates for the unweighted and weighted risk scores are positive and statistically significant ($p>0.01$), supporting the validity of the risk score variable. Columns 3-5 confirm the patterns using cross-sectional measures of BMI. The point estimate for the genetic risk score is significant for all cases ($p<0.01$). However, the SES-health gradients are imprecisely estimated, most likely due to smaller subsamples. The results for the binary logit models (columns 6 and 7) yield similar conclusions: the associations between health outcomes and risk scores are significant in both models (normal weight=0, overweight=1; normal weight=0, obese=1), but the parental education-health gradient is now statistically significant only when the overweight indicator is used as the outcome variable. The LPC only provides data on parental information for one cross-section (1980). Thus, the data limits possibilities for the IV analysis. Notably, there are no measures that could be used to explain family education. As an instrument for family income, we used a variable that indicates the average wage level in the father's industry of occupation. The instrument has a strong first stage ($F=36.4$; see column 8). However, the estimates for parental effects were imprecise and did not differ statistically from the OLS estimates.

5. Discussion

Recent research in health genetics has identified genetic endowments that are robustly associated with health traits (Belsky et al., 2013). This paper uses linked YFS-FLEED-LPC data to provide new evidence for five associations. In particular, we examine whether parental SES moderates these relationships. To accomplish this goal, we exploit genetic risk scores constructed using risk variants identified in GWASs, longitudinal biomarker data obtained from examinations by health-care professionals, and measures of parental resources using administrative data.

The rich data, which combine comprehensive register-based data and objectively measured health information, provided several strengths. First, we examined the relationships between individual genetic backgrounds determined at conception and multiple health outcomes. Health outcomes based on professional examinations conducted at health centres provided accuracy for analyses compared to self-rated surveys, which are vulnerable to non-random measurement or reporting errors. In addition, we measured health outcomes at several points in time, thereby minimizing biases related to short-term variations in health. Second, the measures of parental resources were drawn from official registers. This approach minimizes the errors that plague retrospective questions on participants' childhood environments. The use of two complementary measures for parental background provided a detailed account of parental resources. In particular, education serves as an indicator of longer term SES compared to family income, which is prone to short-term variation. Education may also improve parental skills and knowledge, which contribute to the ability of parents to transmit human capital to their offspring beyond financial resources. Third, the data provided measures for the participants' investments in human capital and for their

initial health endowment. Therefore, our approach accounted both for the early endowment effects and the endowment effects that accumulate throughout childhood. For a comprehensive analysis of the role of parental resources in children's later development, it is important to incorporate both aspects (Black et al., 2007; Cook and Fletcher, 2015).

We reported three important findings using the linked data. First, the risk scores identified in GWA studies contribute to health outcomes. Therefore, the genetic inputs determine a non-significant fraction of health traits across individuals. The associations were robust for the use of controls for participants' initial health endowment, human capital investments, health-related behaviours, and family resources. Second, the parental resources were related to health outcomes, and these associations varied by resource measure and biomarker. In particular, family education was negatively related to obesity (BMI and WHR), whereas family income was negatively related to LDL cholesterol and triglyceride levels. Third, we did not find evidence for statistically significant or economically relevant interactions between parental resources and genetic risk scores although there was evidence that high family income may modify poor genetic background, particularly for the triglyceride levels. The robustness of the results was investigated by estimating the gene-SES gradients for the gender and age-based participant cohorts.

Empirical results regarding the associations of genetic risk scores, environment, and health outcomes with the possible moderation of the genetic effect by environment have policy implications if these associations are quantified with accuracy. In particular, the existence of gene-environment interaction effects would support policy interventions in which target groups are defined by environmental status, such as SES

(Conley et al., 2015). Our results show that both genetic inputs and family resources contribute to health. Therefore, the health outcomes examined in the study can be improved by policies that improve parental resources. However, the SES-health gradients in the YFS data appeared to be weaker than those documented for other developed countries (Currie, 2009; Andersson, 2016). These results may be related to the scope of Finnish public health care, which is comprehensive and inexpensive compared to that of other countries, thus providing universal access to health care across SESs (Walhback et al., 2008). In addition, the Finnish school system encourages healthy behaviour by compulsory education in sports and health (Kannas, 2004). We found no evidence for the view that genetic risks were mitigated or exacerbated by family resources. This finding contrasts with that from prior research (Björklund et al., 2006; Turkheimer et al., 2003; Caspi 2002, 2003), which has indicated that the adverse effects of genetic risk may be modified by parental SES. Intuitively, parental SES could also improve health outcomes more generally. For instance, higher parental income and education can provide material and non-material resources that promote children's health behaviour or means for acquiring health care (Case et al., 2005). There are several possible factors that may explain the negligible modifying effects in the health-gene gradients reported in this study. First, the average age of participants in the sample was 36 years. Therefore, it is possible that the increased genetic risks were not yet fully developed among the participants. For instance, Rietveld et al. (2016) report that genetic risk scores predict educational attainment better in older than in younger cohorts. Second, the results may reflect the measurement of parental resources (family income and family education). Alternative measures could include parental wealth, maternal education, or father's employment (Berchick, 2016).

We suggest three extensions for future research, all of which are related to the limitations of our study. First, information on parental genotype to account for the influence of parental gene variants on parental SES would be of substantial value. This information could provide insight into passive gene-environment correlations, i.e., how parents' genes influence their environments and thus the parental SES (Lundborg and Stenberg, 2010). The analysis could extend research by Conley et al. (2015), who examined the effects of parental education on offspring and found that the only significant interaction effect was between maternal genotype and offspring genotype. Second, the G*E analysis could be extended to environmental exposures that exhibit social gradients. For instance, our analysis could be augmented using measures for parental health behaviour. In addition, because these measures may be correlated with children's health behaviour and thus their later health traits, they are possible determinants of family resources. Together with other refined measures of parental behaviour, this approach could constitute an alternative strategy for the use of instrumental variable methods or for the use of data on parental genotypes. Third, empirical results regarding the associations between individuals' health, genetic risks, and environment can only be used for policy purposes if these associations are quantified with adequate accuracy. This objective requires the identification of the causal effects, as stressed in Lundborg and Stenberg (2010) and Belsky and Israel (2014). In addition, as emphasized in prior studies (Conley, 2016; Conley et al., 2015; Kuenhle, 2014), there is always concern regarding the identification of the interaction term. In the context of observational data, a proper study design is difficult to establish. For instance, the linked YFS data in their current form contain only few alternative instruments for parental SES. In the future, the linked YFS-FLEED-LPC data could be

augmented by prior waves of the Longitudinal Population Census (e.g., for 1975) to provide additional instruments for parental background.

References

Alberti, K. G. M. M., & Zimmet, P. F. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. *Diabetic Medicine*, *15*(7), 539-553.

Allen, H. L., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., ... & Ferreira, T. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, *467*(7317), 832-838.

Andersson, M. A. (2016). Health returns to education by family socioeconomic origins, 1980–2008: Testing the importance of gender, cohort, and age. *SSM-Population Health*, *2*, 549-560.

Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., ... & Poulton, R. (2016). The genetics of success: How single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Psychological Science*, *27*(7), 957-972.

Belsky, D. W., & Israel, S. (2014). Integrating genetics and social science: Genetic risk scores. *Biodemography and Social Biology*, *60*(2), 137-155.

Belsky, D. W., Moffitt, T. E., & Caspi, A. (2013). Genetics in population health science: strategies and opportunities. *American Journal of Public Health*, *103*(S1), S73-S83.

Berchick, E. R. (2016). The relationship between maternal education and reported childhood conditions. *Social Science & Medicine*, *170*, 170-179.

Björklund, A., & Jäntti, M. (2012). How important is family background for labor-economic outcomes?. *Labour Economics*, *19*(4), 465-474.

Björklund, A., & Salvanes, K.G. (2010). Education and family background: mechanism and policies. In Hanushek, E., Machin, S., Woessman, L. (Eds.), Handbooks in Economics and Education, 3 Amsterdam: North-Holland, pp. 201-247.

Björklund, A., Lindahl, M., & Plug, E. (2006). The origins of intergenerational associations: Lessons from Swedish adoption data. *The Quarterly Journal of Economics*, *121*(3), 999-1028.

Black, S.E., & Devereux, P.J. (2010). Recent developments in intergenerational mobility. In Ashelfelter, O., Card, D. (Eds.), Handbook of Labor Economics. Amsterdam: North-Holland, pp. 1487-1541.

Black, S.E., Devereux, P.J., & Salvanes, K.G. (2007). From cradle to the labour market? The effect of birth weight on adult outcomes. *Quarterly Journal of Economics*. 122(1), 409-439.

Boronat, M., Saavedra, P., Varillas, V. F., Wagner, A. M., López-Plasencia, Y., Alberiche, M. P., & Nóvoa, F. J. (2009). Differences in traditional and emerging cardiovascular risk

factors of subjects discordantly classified by metabolic syndrome definitions of the International Diabetes Federation and the National Cholesterol Education Program. *Nutrition, Metabolism and Cardiovascular Diseases*, *19*(6), 417-422.

Bouzigon, E., Corda, E., Aschard, H., Dizier, M. H., Boland, A., Bousquet, J., ... & Scheinmann, P. (2008). Effect of 17q21 variants and smoking exposure in early-onset asthma. *New England Journal of Medicine*, *359*(19), 1985-1994.

Böckerman, P., & Maczulskij, T. (2016). The Education-health Nexus: Fact and fiction. *Social Science & Medicine*, *150*, 112-116.

Böckerman, P., Viinikainen, J., Vainiomäki, J., Hintsanen, M., Pitkänen, N., Lehtimäki, T., ... & Raitakari, O. (2017). Stature and long-term labor market outcomes: Evidence using Mendelian randomization. *Economics and Human Biology*, *24*, 18-29.

Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*, *24*(2), 365-389.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., ... & Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*, *301*(5631), 386-389.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., ... & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, *297*(5582), 851-854.

Conley, D. (2016). Socio-Genomic Research Using Genome-Wide Molecular Data. *Annual Review of Sociology*, *42*, 275-299.

Conley, D., Domingue, B. W., Cesarini, D., Dawes, C., Rietveld, C. A., & Boardman, J. D. (2015). Is the effect of parental education on offspring biased or moderated by genotype?. *Sociological Science*, *2*, 82-105.

Conti, G., & Heckman, J. J. (2010). Understanding the early origins of the education–health gradient: A framework that can also be applied to analyse gene–environment interactions. *Perspectives on Psychological Science*, *5*(5), 585-605.

Cook, C. J., & Fletcher, J. M. (2015). Understanding heterogeneity in the effects of birth weight on adult cognition and wages. *Journal of Health Economics*, *41*, 107-116.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, *47*(1), 87-122.

Foley, D. L., Eaves, L. J., Wormley, B., Silberg, J. L., Maes, H. H., Kuhn, J., & Riley, B. (2004). Childhood Adversity, Monoamine Oxidase A Genotype, and Risk for ConductDisorder. *Archives of General Psychiatry*, *61*(7), 738-744.

Gatto, N. M., Campbell, U. B., Rundle, A. G., & Ahsan, H. (2004). Further development of the case-only design for assessing gene–environment interaction: evaluation of and adjustment for bias. *International Journal of Epidemiology*, *33*(5), 1014-1024.

Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Steinthorsdottir, V., ... & Workalemahu, T. (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics*, *42*(11), 949-960.

Hernesniemi, J. A., Lyytikäinen, L. P., Oksala, N., Seppälä, I., Kleber, M. E., Mononen, N., ... & Martiskainen, M. (2015). Predicting sudden cardiac death using common genetic risk variants for coronary artery disease. *European Heart Journal*, 34(26), 1669-1675.

Hoffjan, S., Nicolae, D., Ostrovnaya, I., Roberg, K., Evans, M., Mirel, D. B., ... & Gern, J. E. (2005). Gene-environment interaction effects on the development of immune responses in the 1st year of life. *The American Journal of Human Genetics*, *76*(4), 696-704.

Hunter, D. J. (2005). Gene–environment interactions in human diseases. *Nature Reviews Genetics*, *6*(4), 287-298.

Kannas L. 2004. School children's health and health behaviour in change. HBSC study 20 years. University of Jyväskylä, Research Centre for Health Promotion: Jyväskylä.

Kilpeläinen, T. O., Qi, L., Brage, S., Sharp, S. J., Sonestedt, E., Demerath, E., ... & Holzapfel, C. (2011). Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PloS Med*, *8*(11), e1001116.

Kuehnle, D. (2014). The causal effect of family income on child health in the UK. *Journal of Health Economics*, *36*, 137-150.

Kunst, A. E., Bos, V., Lahelma, E., Bartley, M., Lissau, I., Regidor, E., ... & Helmert, U. (2005). Trends in socioeconomic inequalities in self-assessed health in 10 European countries. *International Journal of Epidemiology*, *34*(2), 295-305.

Lechner, M. (2009). Long-run labour market and health effects of individual sports activities. *Journal of Health Economics*, *28*(4), 839-854.

Lundborg, P., & Stenberg, A. (2010). Nature, nurture and socioeconomic policy—What can we learn from molecular genetics?. *Economics & Human Biology*, *8*(3), 320-330.

Mackenbach, J. P., Kunst, A. E., Cavelaars, A. E., Groenhof, F., Geurts, J. J., & EU Working Group on Socioeconomic Inequalities in Health. (1997). Socioeconomic inequalities in morbidity and mortality in western Europe. *The Lancet*, *349*(9066), 1655-1659.

Mansikkaniemi, K., Juonala, M., Taimela, S., Hirvensalo, M., Telama, R., Huupponen, R., ... & Jula, A. (2012). Cross-sectional associations between physical activity and selected coronary heart disease risk factors in young adults. The Cardiovascular Risk in Young Finns Study. *Annals of medicine*, *44*(7), 733-744.

Mattei, J., Qi, Q., Hu, F. B., Sacks, F. M., & Qi, L. (2012). TCF7L2 genetic variants modulate the effect of dietary fat intake on changes in body composition during a weight-loss intervention. *The American journal of clinical nutrition*, *96*(5), 1129-1136.

McLaren, L. (2007). Socioeconomic status and obesity. *Epidemiologic Reviews*, *29*(1), 29-48.

Mendis, S., Puska, P., & Norrving, B. (2011). *Global atlas on cardiovascular disease prevention and control*. World Health Organization.

National Institute for Health and Welfare. (2014). Perinatal statistics: parturients, deliveries and newborns in 2013 – Statistical report. Retrieved http://www.julkari.fi/bitstream/ handle/10024/ 116818/Tr23_14.pdf (accessed 7.3.2017)

Palmer, T. M., Lawlor, D. A., Harbord, R. M., Sheehan, N. A., Tobias, J. H., Timpson, N. J., ... & Sterne, J. A. (2012). Using multiple genetic variants as instrumental variables for modifiable risk factors. *Statistical methods in medical research*, *21*(3), 223-242.

Raitakari, O. T., Juonala, M., Kähönen, M., Taittonen, L., Laitinen, T., Mäki-Torkko, N., ... & Åkerblom, H. K. (2003). Cardiovascular risk factors in childhood and carotid artery intima-media thickness in adulthood: the Cardiovascular Risk in Young Finns Study. *Jama*, *290*(17), 2277-2283.

Raitakari, O. T., Juonala, M., Rönnemaa, T., Keltikangas-Järvinen, L., Räsänen, L., Pietikäinen, M., ... & Jula, A. (2008). Cohort profile: the cardiovascular risk in Young Finns Study. *International Journal of Epidemiology*, *37*(6), 1220-1226.

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., ... & Albrecht, E. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, *340*(6139), 1467-1471.

Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., ... & Randall, J. C. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics*, *42*(11), 937-948.

Tarkiainen, L., Martikainen, P., Laaksonen, M., & Valkonen T. (2012). Trends in life expectancy by income from 1988 to 2007: decomposition by age and cause of death. *Journal of Epidemiol Community*, 66, 573-578.

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., ... & Johansen, C. T. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, *466*(7307), 707-713.

Thompson, O. (2014). Economic Background and Educational Attainment The Role of Gene-Environment Interactions. *Journal of Human Resources*, *49*(2), 263-294.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, *14*(6), 623-628.

Valkonen, T., Sihvonen, A. P., & Lahelma, E. (1997). Health expectancy by level of education in Finland. *Social Science & Medicine*, *44*(6), 801-808.

Wahlbeck, K., Manderbacka, K., Vuorenkoski, L., Kuusio, H., Luoma, M. L., & Widström, E. (2008). Quality and equality of access to healthcare services. *HealthQUEST country report for Finland. Helsinki: National Research and Development Centre for Welfare and Health*.

World Health Organization. 2017a. Obesity and overweight. Retrieved http://www.who.int/mediacentre /factsheets /fs311/en/ (accessed 10.3.2017).

World Health Organization. 2017b. Guidelines for the management of dyslipidaemia in patients with diabetes mellitus. Retrieved http://applications.emro.who.int/dsaf/dsa699.pdf (accessed 10.3. 2017).

Table 1 Descriptive statistics; full sample and sample averages by income and education groups.

	Full sample	**Samples by family income**				**Samples by family education**			
		Family income low	Family income high	Difference	t-statistics	Family education low	Family education high	Difference	t-statistics
Biomarker									
BMI	25.83 (4.60)	26.12 (4.70)	25.53 (4.49)	0.591	3.05***	25.99 (4.62)	24.80 (4.36)	1.187	4.19***
WHR	0.87 (0.08)	0.87 (0.08)	0.86 (0.08)	0.010	2.89***	0.87 (0.08)	0.85 (0.082)	0.018	3.63***
Triglycerides	1.36 (0.80)	1.38 (0.80)	1.33 (0.81)	0.048	1.43	1.36 (0.81)	1.30 (0.73)	0.060	1.324
HDL cholesterol	1.31 (0.30)	1.29 (0.30)	1.33 (0.30)	-0.034	-2.62***	1.31 (0.30)	1.33 (0.31)	-0.024	-1.23
LDL cholesterol	3.19 (0.73)	3.22 (0.74)	3.16 (0.73)	0.061	1.97**	3.19 (0.74)	3.18 (0.73)	0.007	0.15
Genetic risk score, unweighted (GRS)									
BMI	29.13 (3.35)	29.25 (3.33)	29.02 (3.36)	0.231	1.64	29.11 (3.34)	29.29 (3.40)	-0.178	-0.86
WHR	16.26 (2.52)	16.20 (2.55)	16.31 (2.48)	-0.114	-1.07	16.25 (2.53)	16.31 (2.40)	-0.065	-0.41
Triglycerides	0.99 (0.09)	0.99 (0.10)	0.98 (0.09)	0.006	1.64	0.99 (0.09)	0.99 (0.09)	0.001	0.10
HDL cholesterol	44.68 (3.69)	44.62 (3.81)	44.73 (3.57)	-0.109	-0.68	44.68 (3.73)	44.67 (3.44)	0.011	0.05
LDL cholesterol	0.95 (0.07)	0.95 (0.08)	0.95 (0.07)	0.000	0.04	0.95 (0.07)	0.96 (0.073)	-0.008	-1.68*
Parental SES (1980)									
Log of family income	9.31 (0.84)	8.84 (0.95)	9.78 (0.28)	-0.938	-32.09***	9.22 (0.82)	9.90 (0.71)	-0.678	-15.090***
High education	0.13 (0.34)	0.03 (0.17)	0.24 (0.43)	-0.208	-15.40***	-	-	-	-
Participant characteristics									
Average age, 2010	40.37 (5.01)	40.07 (5.04)	40.67 (4.96)	-0.596	-2.85***	40.52 (4.99)	39.37 (5.03)	1.157	3.763***
Female, share	0.54 (0.50)	0.53 (0.50)	0.54 (0.50)	-0.004	-0.171	0.54 (0.50)	0.51 (0.50)	0.027	0.881
Education, 2010	0.37 (0.48)	0.30 (0.46)	0.43 (0.50)	-0.132	-5.48***	0.32 (0.47)	0.66 (0.48)	-0.334	-9.711***
Log of income, 2001	8.62 (3.00)	8.47 (3.10)	8.77 (2.90)	-0.295	-1.945*	8.65 (2.97)	8.45 (3.21)	0.205	0.928
Birth weight (grams)	3498.9(545.3)	3528.7(521.8)	3470.1(566.0)	58.594	2.132**	3501.4(540.6)	3483.4(575.0)	17.998	0.449
Fruit and vegetable consumption at age 15	2.93 (0.87)	2.81 (0.86)	3.04 (0.85)	-0.229	-5.288***	2.88 (0.86)	3.23 (0.83)	-0.354	-5.621***
Sports intensity at age 15	2.16 (0.53)	2.11 (0.52)	2.20 (0.53)	-0.090	-3.383***	2.14 (0.52)	2.27 (0.57)	-0.132	-3.220***
Sports frequency at age 15	4.97 (1.57)	4.86 (1.61)	5.08 (1.52)	-0.220	-2.780***	4.95 (1.59)	5.13 (1.43)	-0.189	-1.768*

Notes: The table values represent the mean and standard deviation (in parentheses). Biomarkers are measured as the average value of the measurements in 2001, 2007 and 2011. In the case of missing values, the average is calculated based on two or one year values. *significant at the 0.10 level; **at the 0.05 level; *** at the 0.01 level.

Table 2 Associations between health outcome, genetic risk score, and parental resources.

	HDL			LDL			Triglycerides			BMI			WHR		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Risk score (GRS)	0.011*** (0.001)	0.011*** (0.001)	0.011*** (0.001)	0.727*** (0.064)	0.733*** (0.063)	0.791*** (0.071)	0.905*** (0.096)	0.911*** (0.096)	0.852*** (0.114)	0.007*** (0.001)	0.007*** (0.001)	0.008*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002** (0.001)
Parental SES															
***lncome (SES1)**		0.008 (0.005)	0.008 (0.006)		-0.023*** (0.005)	-0.020*** (0.006)		-0.025** (0.012)	-0.020 (0.013)		-0.005 (0.004)	-0.000 (0.005)		-0.004* (0.002)	-0.002 (0.002)
***Education (SES2)**		0.017 (0.014)	0.008 (0.016)		0.015 (0.013)	0.027 (0.017)		-0.028 (0.029)	-0.021 (0.034)		-0.040*** (0.010)	-0.033*** (0.012)		-0.015*** (0.005)	-0.011** (0.005)
Covariates															
***Education**			-0.006 (0.012)			-0.017 (0.012)			-0.015 (0.025)			-0.034*** (0.009)			-0.014*** (0.004)
***lncome**			0.005*** (0.002)			-0.001 (0.002)			0.002 (0.003)			-0.001 (0.002)			-0.001 (0.001)
***Diet**			0.002 (0.006)			-0.018*** (0.006)			-0.022* (0.013)			-0.010*** (0.005)			-0.003 (0.002)
***Sports intensity**			0.031*** (0.011)			-0.006 (0.011)			-0.053*** (0.021)			-0.004 (0.008)			-0.008** (0.003)
***Sports frequency**			-0.001 (0.004)			0.004 (0.004)			0.002 (0.008)			0.004 (0.003)			0.006 (0.001)
***Birth weight**			0.000 (0.000)			0.000 (0.000)			-0.000*** (0.000)			-0.000*** (0.000)			-0.000 (0.000)
F-test (p-value) for SES	-	2.50 (0.082)	1.05 (0.349)	-	9.04 (0.000)	5.49 (0.004)	-	3.83 (0.022)	1.56 (0.211)	-	9.97 (0.000)	4.28 (0.014)	-	9.31 (0.000)	3.02 (0.049)
Incremental R2, %-Points	3.1	0.2	4.2	5.3	0.6	3.3	3.3	0.3	1.9	1.8	0.7	3.2	0.2	0.4	1.0
x-stand. GRS coefficient	3.9	3.9	3.9	5.6	5.6	5.6	9.3	9.3	9.3	2.3	2.3	2.3	0.5	0.5	0.5
Obs.	2,167	2,167	1,488	2,282	2,282	1,558	2,294	2,294	1,567	2,266	2,266	1,553	2,269	2,269	1,552

Notes: Unless otherwise stated, the tabled values are unstandardized regression coefficients and heteroscedasticity-robust standard errors (in parentheses). Biomarkers are measured as the average value of the measurements in 2001, 2007 and 2011. In the case of missing values, the average is calculated based on two- or one-year values**.** All models include predetermined covariates for the birth month, birth year effects, gender, native language and region of birth. Participants' own education and income were measured in 2010 and 2001, respectively, as well as their fruit and vegetable consumption and sports frequency and intensity at the age of 15. Parental background information refers to year 1980. The dependent variables, parental income and participants' own income were ln-transformed because of skewed distributions; all the other variables are in levels. *significant at the 0.10 level; **at the 0.05 level; *** at the 0.01 level. Baseline for incremental R2 includes only predetermined covariates without GRS.

Table 3 Gene-health gradient and parental resources.

	Family income (SES1)					**Family Education (SES2)**				
	Full sample	Females only	Males only	Young cohorts	Old cohorts	Full sample	Females only	Males only	Young cohorts	Old cohorts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Health outcome										
BMI	-0.045 (0.033)	-0.113** (.048)	0.056 (0.045)	-0.081 (0.065)	-0.015 (0.036)	-0.024 (0.091)	-0.039 (0.148)	-0.015 (0.101)	-0.145 (0.105)	0.104 (0.158)
Obs.	2,266	1,215	1,051	1,154	1,112	2,266	1,215	1,051	1,154	1,112
WHR	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.001)	0.001 (0.001)	-0.001 (0.001)	-0.001 (0.002)	-0.003 (0.002)	0.003 (0.002)	-0.000 (0.002)	-0.003 (0.003)
Obs.	2,269	1,216	1,053	1,157	1,112	2,269	1,216	1,053	1,157	1,112
Triglycerides	-0.447*** (0.163)	-0.321** (0.135)	-0.545 (0.367)	-0.331* (0.195)	-0.517* (0.293)	-0.185 (0.429)	0.478 (0.466)	-0.775 (0.775)	0.165 (0.572)	-0.882 (0.610)
Obs.	2,294	1,231	1,063	1,167	1,127	2,294	1,231	1,063	1,167	1,127
HDL	0.004* (0.002)	0.002 (0.003)	0.007** (0.003)	0.002 (0.003)	0.004* (0.002)	-0.008 (0.006)	-0.010 (0.009)	-0.005 (0.008)	-0.010 (0.007)	-0.007 (0.010)
Obs.	2,167	1,162	1,005	1,107	1,060	2,167	1,162	1,005	1,107	1,060
LDL	-0.051 (0.191)	-0.335 (0.240)	0.330 (0.232)	-0.151 (0.228)	0.149 (0.369)	-0.165 (0.558)	-0.167 (0.672)	0.157 (0.957)	-0.437 (0.649)	0.412 (1.114)
Obs.	2,282	1,230	1,052	1,163	1,119	2,282	1,230	1,052	1,163	1,119

Notes: The tabled values are unstandardized regression coefficients for the biomarker-SES interaction terms and heteroscedasticity-robust standard errors (in parentheses). Biomarkers are measured as the average value of the measurements in 2001, 2007 and 2011. In the case of missing values, the average is calculated based on two- or one-year values**.** All models include predetermined covariates for the birth month, birth year effects, gender, native language and region of birth. Parental background information refers to year 1980. The dependent variables and parental income were ln-transformed because of skewed distributions; all the other variables are in levels. *significant at the 0.10 level; **at the 0.05 level; *** at the 0.01 level. The age cohorts “young” and “old” refer to individuals born 1971-1977 (the three youngest age cohorts) and 1962-1968 (the three oldest age cohorts), respectively.

Table 4 Model extensions and robustness; results for BMI.

	Genetic risk score		**Cross-sectional results**			**Logit**		**IV**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Unweighted	Weighted	2001	2007	2011	Overweight	Obese	
Risk score (GRS)	0.008*** (0.001)	0.052*** (0.008)	0.007*** (0.001)	0.007*** (0.001)	0.009*** (0.002)	0.018*** (0.004)	0.012*** (0.003)	0.008*** (0.001)
Parental SES								
***Income (SES1)**	-0.000 (0.005)	-0.000 (0.005)	0.001 (0.005)	-0.004 (0.006)	-0.005 (0.009)	-0.001 (0.016)	-0.005 (0.008)	-0.006 (0.034)
***Education (SES 2)**	-0.033*** (0.012)	-0.033*** (0.012)	-0.024* (0.013)	-0.029** (0.013)	-0.030** (0.015)	-0.095** (0.038)	-0.043 (0.032)	-
Covariates								
***Education**	-0.034*** (0.009)	-0.034*** (0.009)	-0.033*** (0.010)	-0.037*** (0.010)	-0.039*** (0.011)	-0.115*** (0.026)	-0.077*** (0.022)	-0.035*** (0.013)
***Income**	-0.001 (0.002)	-0.001 (0.002)	-0.003 (0.002)	0.001 (0.002)	-0.001 (0.002)	-0.006 (0.004)	-0.001 (0.003)	-0.001 (0.002)
***Diet**	-0.010** (0.005)	-0.009* (0.005)	-0.011** (0.005)	-0.009* (0.005)	-0.013** (0.006)	-0.028** (0.014)	-0.009 (0.011)	-0.009* (0.005
***Activity (intensity)**	-0.004 (0.008)	-0.004 (0.008)	0.001 (0.009)	-0.008 (0.009)	-0.007 (0.009)	0.003 (0.024)	-0.023 (0.018)	-0.006 (0.009)
***Activity (frequency)**	0.004 (0.003)	0.004 (0.003)	0.008** (0.003)	0.003 (0.003)	0.003 (0.004)	0.013 (0.009)	0.007 (0.007)	0.004 (0.003)
***Birth weight**	0.000*** (0.000)	0.000*** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)	0.000* (0.000)	0.000*** (0.000)
F-test for SES (p-value) /Chi2 (p-value)	4.28 (0.014)	4.21 (0.015)	1.84 (0.160)	3.12 (0.045)	2.88 (0.057)	6.54 (0.038)	2.70 (0.259)	
First stage F-statistics								36.26
R2 / Pseudo R2	0.117	0.116	0.113	0.110	0.113	0.070	0.070	-
N	1,553	1,553	1,344	1,272	1,116	1,553	1,553	1,409

Notes: The tabled values are unstandardized regression coefficients and heteroscedasticity-robust standard errors (in parentheses). In columns (6) and (7) the tabled values are average marginal effects and standard errors (in parentheses). BMI is measured as the average value of the measurements in 2001, 2007 and 2011. In the case of missing values, the average is calculated based on two- or one year-values**.** All models include the following covariates: birth month, birth year, gender, native language, region of birth, birth weight, participants' own education and income (in 2010 and 2001, respectively), fruit and vegetable consumption (at age 15) and sports frequency and intensity (at age 15). Parental background information refers to year 1980. The dependent variables, parental income and participants' own income were ln-transformed because of skewed distributions; all the other variables are in levels. *significant at the 0.10 level; **at the 0.05 level; *** at the 0.01 level.

Appendix 1

Figure A1: Risk score distributions by family education (university-level versus no)

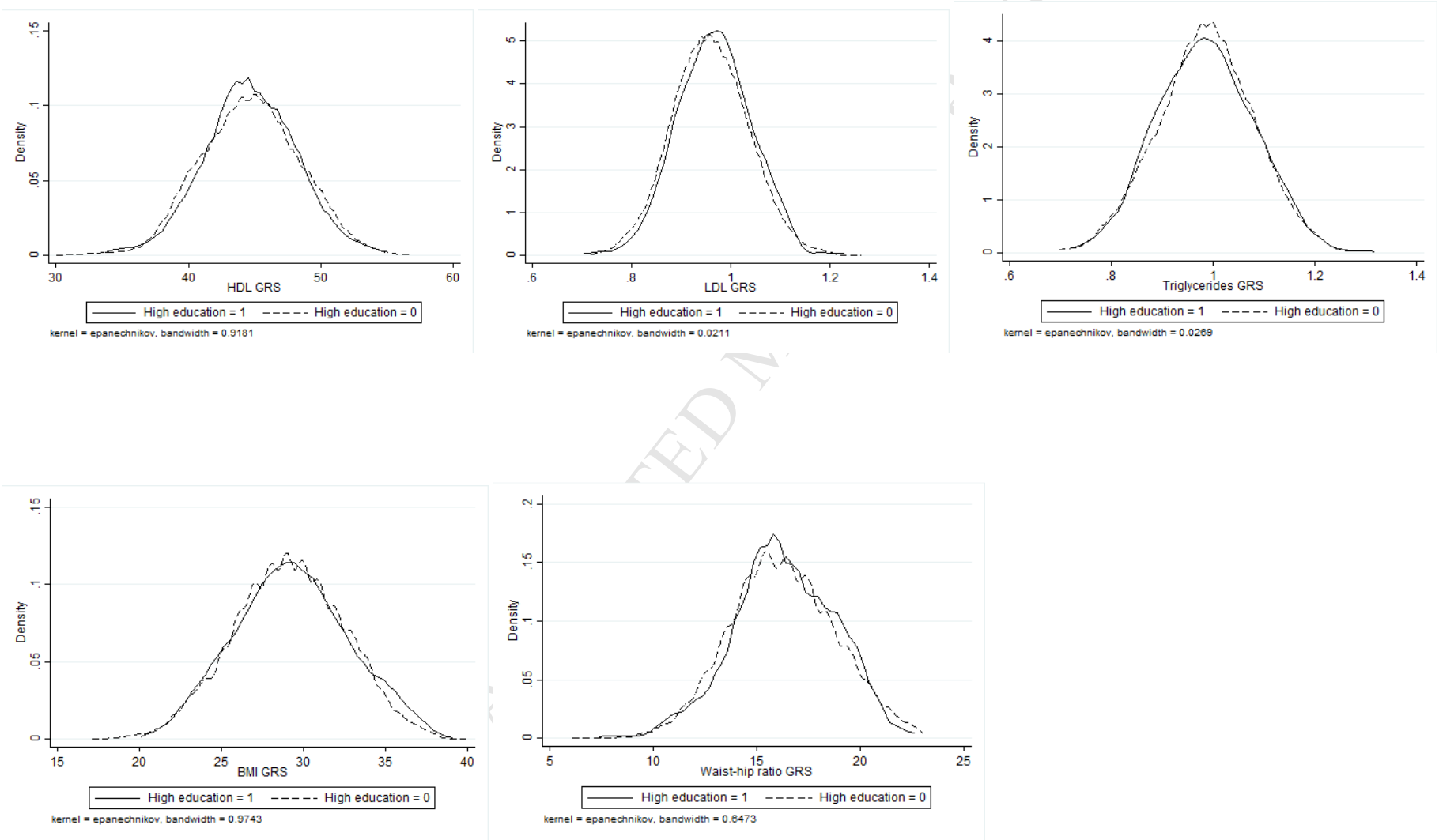
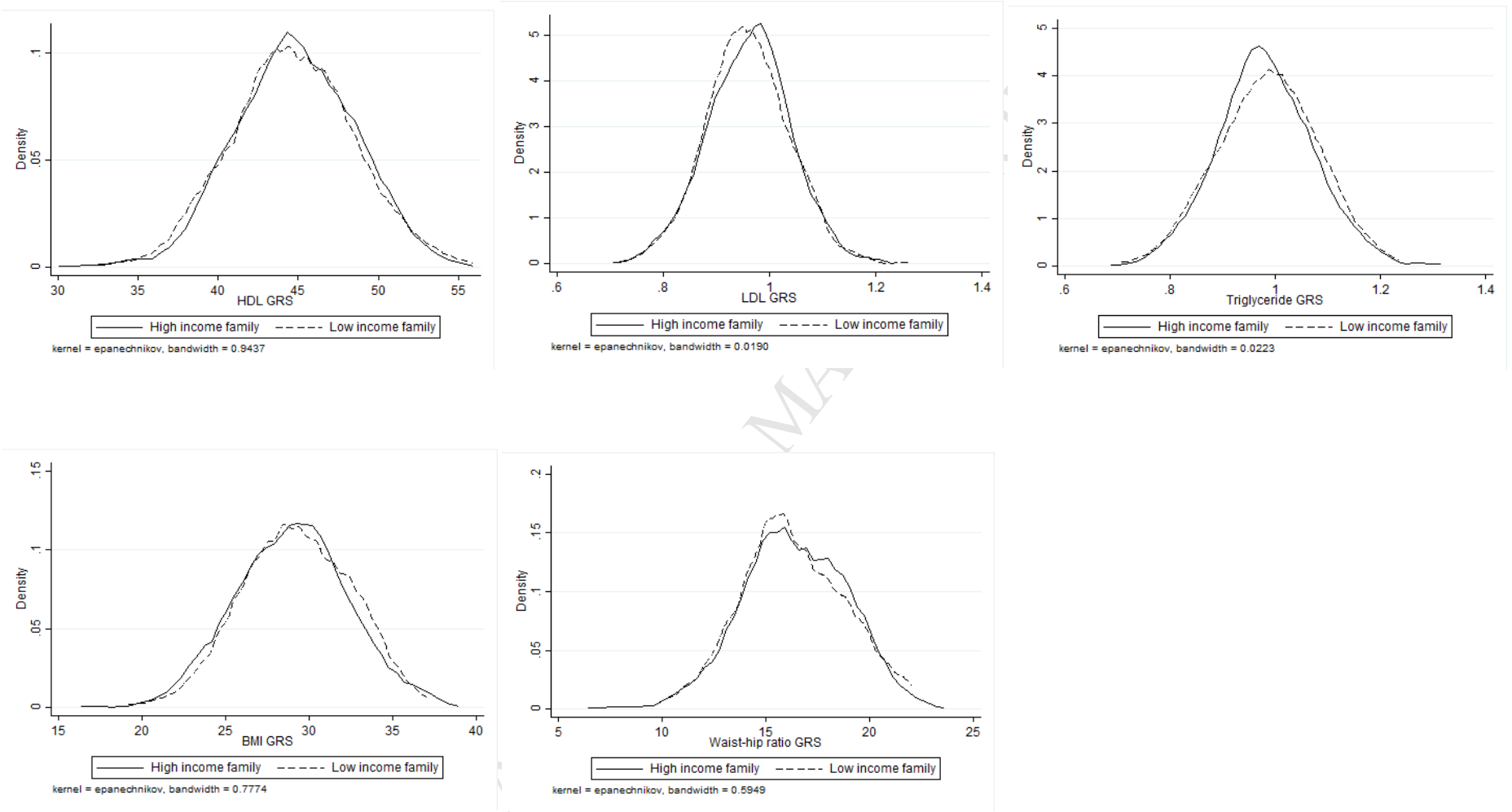


Figure A2: Risk score distributions by family income (highest and lowest quartiles)



Acknowledgements

The Young Finns Study has been financially supported by the Academy of Finland: grants 286284 (for T.L.), 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); 293120 (Strategic Research Council funding for the project Work, Inequality and Public Policy); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (for T.L. X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research and Finnish Cultural Foundation; Tampere Tuberculosis Foundation and Emil Aaltonen Foundation. The Palkansaaja Foundation supported the use of linked data. Jutta Viinikainen and Jaakko Pehkonen acknowledge financial support from the Yrjö Jahnsson Foundation (grants 6664 and 6646).

Highlights

Highlights:

Genetic risk scores have an independent effect on adult health.

Genetic risk is not mitigated or amplified by parental resources.

Parental resources are related to adult health outcomes.

Family education is negatively related to obesity and the waist-hip ratio.

Family income is negatively related to LDL cholesterol and triglyceride levels.