Nature and nurture in the production of childhood BMI *

Uta Bolt,† Gabriella Conti‡, Nancy Daza Baez‡ and Rong Hai§

January 31, 2025

PRELIMINARY AND INCOMPLETE

Abstract

This paper estimates the first BMI production function which incorporates both child and parental genetic and environmental inputs into a dynamic framework. We investigate the relative impact of genetic endowment and parents' BMI on a child's BMI and related behaviors. Our data come from the UK Millennium Cohort Study, a longitudinal study that follows children from ages 3 to 17, and includes BMI and genetic information for both children and parents. We find that the genetic predisposition exerts a cumulative influence on BMI as children grow older from age 3 to age 17. We also find that a significant extent of the differences in child BMI, dietary behaviors, and exercise can be explained by parents' BMI. Moreover, while maternal BMI affects both dietary behaviors and exercise, paternal BMI primarily influences exercise. These findings highlight the strong influence of parents' BMI on children's health beyond genetic factors, and suggest that targeted early interventions focusing on family-based dietary and exercise habits may offer effective strategies for addressing childhood obesity.

Keywords: BMI, Obesity, Genetic Endowment JEL codes: I12, Q56

*We gratefully acknowledge funding from the the European Commission through the Horizon 2020 research and innovation programme under grant agreements no. 874739 (LongITools) and no. 819752 (DEVORHBIOSHIP).

†University of Bristol

‡University College London

§University of Miami

1 Introduction

The escalating global prevalence of childhood obesity, a condition with significant consequences for long-term health and socioeconomic outcomes (National Health System, 2024; World Health Organization, 2024), has ignited discussions about the critical interplay between genetic predisposition ("nature") and environmental or behavioral influences ("nurture") in shaping metabolic health. Previous research have already established clear connections between obesity, genetic risk, and socioeconomic factors (Fan et al., 2023; Silventoinen and Konttinen, 2020; Yang et al., 2007); however, substantial gaps remain in our understanding of how these factors interact dynamically across developmental stages. The economics literature has been slow to integrate genetic data into models that effectively separate inherited risks from modifiable inputs such as diet, physical activity, and parental behaviors (Biroli, 2015); this gap significantly hinders the formulation of evidence-based policy interventions. This study addresses this gap by leveraging comprehensive longitudinal data from the UK to investigate how genetic endowment and parental BMI collectively influence a child's BMI trajectory and health-related behaviors from early childhood through adolescence. To the best of our knowledge this is the first estimate production functions of BMI and obesity that incorporate both child's and parental genes to identify the independent effect of child's genes from that of her family.

We utilize data from the Millennium cohort study (MCS), a cohort that was surveyed regularly from birth to adolescence (age 17). In particular, the data contain genetic risk scores for BMI for the cohort member and their parents, and measures of diet and exercise. We link this data set to information on food prices and distances to facilities conducive to exercise (parks, playgrounds etc.), which we use as instruments to estimate the causal effect of investments on BMI. The availability of genetic data on the both the child and the parents allows us to identify the causal effect of genes on BMI.

We first document that the association between BMI polygenic scores (PGS) and BMI strengthens as children grow older. At age 7, a standard deviation increase in BMI PGS is associated with a BMI increase of 0.5 points, while at age 17, this increases to 1.4 points. A higher BMI PGS is also associated with worse diet quality and lower exercise frequency. The persistency of BMI as well as the effect of BMI PGS on diet and exercise amplify the effect of BMI PGS on BMI over age.

Next, we investigate the effect of parents' BMI on child BMI and related behaviors. We find that while genetic predisposition significantly impacts BMI, a notable portion of BMI variation is attributed to parents' BMI. In particular, both maternal and paternal BMI significantly predict a child's BMI, even when controlling for the child's own genetic predisposition. Moreover, maternal BMI influences both diet and exercise, whereas paternal BMI predominantly affects exercise behavior. These results underscore the dual role of parents as conduits of both genetic risk and modifiable environmental factors, challenging

deterministic narratives of obesity and highlighting actionable levers for intervention. By elucidating how nature and nurture coalesce across developmental stages, this study provides critical insights for policies aimed at breaking the intergenerational cycle of metabolic disease.

The direct genetic effect of BMI PGS on BMI is stable over time, but its influence amplifies through the persistence of BMI across ages.

We contribute to the literature on genetics and obesity by studying genetic effects in an integrated framework, allowing us to considered multiple potential mechanism simultaneously. We show that genetic predisposition exerts a cumulative, age-dependent effect on BMI, with its influence amplifying nearly threefold between early childhood and adolescence. In addition, our approach allow us to disentangles the environmental and behavioral pathways through which parental BMI -beyond genetic transmission- shapes child health outcomes, revealing distinct maternal and paternal roles: maternal BMI impacts both dietary habits and physical activity, while paternal BMI primarily affects exercise.

In the following, we first describe the existing literature. In section 3, we turn to describing our data and key facts from our data set. Next, in section 4, we outline our empirical strategy. Section 5 describes our results and section 6 concludes.

2 Literature

The interplay between genetic predisposition and environmental influences in shaping childhood and adolescence metabolic health is a focal point for the understanding of intergenerational health transmission. While epidemiological/medical studies have long emphasized cross-sectional correlations between obesity, genetics, and socioeconomic factors, the economics literature on this topic is surprisingly limited (Biroli, 2015; Fan et al., 2023). Critical gaps persist particularly in integrating genetic data into meaningful structural models to disentangle these forces.

A body of research leverages BMI PGS to disentangle genetic and environmental pathways (Amin et al., 2020; Brittain et al., 2024; Fan et al., 2023). Brittain et al. (2024), for instance, demonstrated that individuals with high genetic risk require significantly greater physical activity -over 2000 additional daily steps- to mitigate obesity risk, highlighting the moderating role of lifestyle choices (e.g. physical activity). Similarly, Fan et al. (2023) find that childhood obesity decreases significative educational outcomes -e.g. a 33% reduction in high school GPA, 14.8% reduction in years of schooling-, though such studies often treat PGS as static instruments rather than dynamic inputs. The use of genetic risk scores as instruments is considered problematic due to pleiotropic effects, where genetic predictors may influence multiple outcomes violating IV assumptions (Cawley et al., 2011; Fang, 2014; VanderWeele et al., 2014). In contrast to this approach we consider childs' PGS as structural input of BMI trajectories that moderate

environmental effects.

Central to this debate is the dual role of parental BMI, which operates as both a genetic proxy and a marker of environmental influence. Twin studies estimate that $40 - 60\%$ of BMI heritability stems from shared genetics (Silventoinen and Konttinen, 2020). However, Cawley et al. (2019) found no evidence of "genetic nurture" effects, suggesting parental influence manifests through behavioral channels - such as dietary habits or sedentary norms - rather than shared genetic risk. We contribute to this literature by using variation in prices and distances to exercise facilities as a measure of parental investment, we causally estimate the impact of diet and exercise on BMI for a large cohort of children at different ages from childhood to adolescence.

Longitudinal studies further reveal that BMI disparities emerge early and widen with age. Libuy et al. (2021) documented how income-based gaps in obesity risk grow by 56% for boys and 42% for girls between ages 5 and 17, driven by diverging access to nutritious foods and recreational opportunities. Bairoliya et al. (2024) emphasized the persistence of early-life conditions, with adolescent BMI explaining a third of later-life disparities. Yet these studies rarely integrate genetic data, leaving unresolved how nature and nurture interact dynamically. Bidirectional pathways compound this complexity: mental health challenges, such as depression, can exacerbate obesity through emotional eating, while obesity itself may worsen mental health via social stigma (Willage, 2018).

Various studies have established that genetic predispositions tend to interact with the environment. Houmark et al. (2024) exploit PGS to estimate human capital production functions of cognitive skills. They document three substantive mechanisms through which genes are linked to child cognitive skills: the direct effect of child genes on skills, the indirect effect of child genes via parental investments, and family genetic influences captured by parental genes. Similarly, Muslimova et al. (2020) found that genetic potential for education interacts with birth order, underscoring how environmental contexts modulate genetic effects. Using a similar approach, Biroli (2015) laid the groundwork by modeling BMI trajectories as functions of genetic heterogeneity and health investments, revealing that individuals with obesity-linked genetic variants face higher costs of adopting healthy behaviors, such as exercise or balanced diets. This insight challenges the efficiency of uniform public health policies and underscores the need for targeted interventions. However, the study did not account for parental genetics, which could help isolate the independent effects of an individual's genes from those of their family.

While prior work has separately examined genetic and parental influences, few studies integrate both while differentiating maternal and paternal effects. Our analysis addresses this gap by leveraging the MCS's unique genetic and longitudinal data to provide estimates of the dynamic production of BMI, including the causal impact of investments and genes, to disentangle how parental BMI - distinct from

shared genetics - shapes children's dietary and exercise behaviors. This approach advances understanding of nurture's role in modulating genetic predispositions across developmental stages.

3 Data

3.1 Millennium Cohort Study

The data we use comes from the Millennium cohort study (MCS), which started with over 19,000 children born in the UK between June 2001 and January 2003. Individuals were first surveyed shortly after birth and then followed up at ages 3, 5, 7, 11, 14, and 17. The data set is uniquely suited for our study for several reasons: First, BMI of the cohort member was measured at every wave. Whilst many papers encounter measurement error issues due to self-reported BMI, in the MCS, BMI was measured and reported by the interviewer, minimizing such concerns. Self-reported parental BMI was also recorded during the age age 3, 5, 7, and 11 waves. Second, the survey contains genetic information on the mother, father and cohort member, allowing us to identify the causal effect of genes in BMI formation. Lastly, measures regarding the diet and exercise frequency of cohort members was reported from age 5 onwards, allowing us to understand the role of health investments.

BMI, Diet, and Exercise Measures BMI of the cohort member was calculated as $BMI = \frac{\text{weight in kg}}{\text{height in m}^2}$. Whilst obesity in adulthood is defined as having a BMI of 30 or higher, in childhood, cutoffs are age- and gender-dependent. CLS uses cutoffs determined by the Extended International (IOTF) Body Mass Index Cut-Offs for Thinness, Overweight and Obesity; which we report in Appendix B.

Diet quality in the MCS was captured through several items: number of portions of fruit consumed per day (all waves), frequency of eating breakfast (all waves), whether child consumes healthy (fruit, nuts, vegetables, yoghurt) or unhealthy snacks (crisps, sweets) (age 5 & 7), frequency of consuming sweetened or artificially sweetened drinks (age 11) , frequency of eating vegetables (age 14 & 17), frequency of consuming fast-food (age 14 & 17). To summarize the information in a single index and minimize measurement error, we estimate a factor model using the available diet quality information at each age and then predict factor scores using the Bartlett scoring method. The Bartlett scoring method reduces measurement error by giving a large weight to measure that contain little noise and a smaller weight to measures that are noisy.

Exercise is measured in each wave from age 5 onwards. The variable captures the number of days an individual takes part in exercise. It is important to note, however, that before age 14, this measure was reported by the mother of the cohort member, whereas at ages 14 and 17, it was self-reported by the cohort member. This means that the measure albeit similar, is not comparable over time.

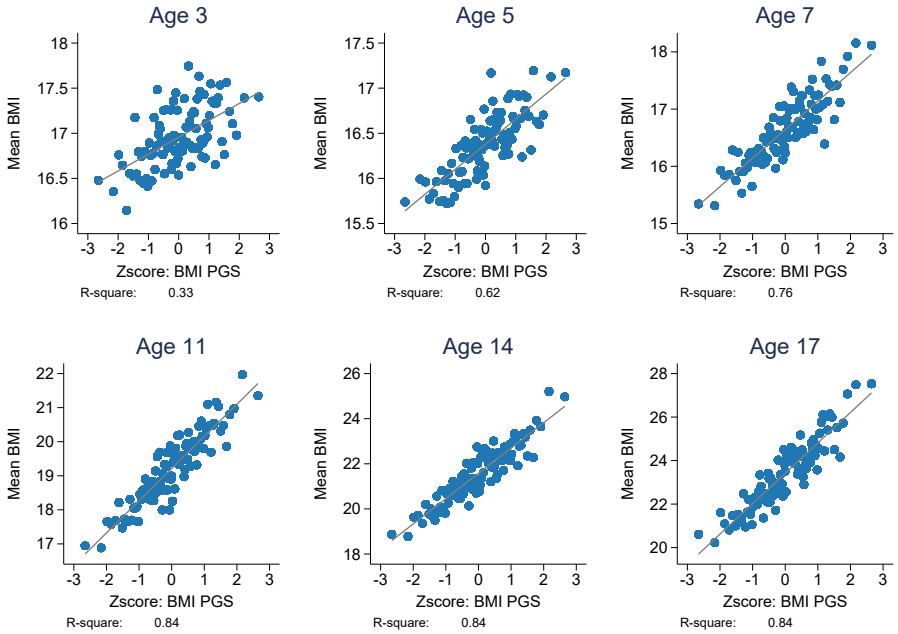
Polygenic Scores To capture individuals' predispositions for BMI, we use polygenic scores. The human genome consists of 23 pairs of chromosomes, one of which is inherited from the mother, the other inherited from the father. These chromosomes are made from DNA (deoxyribonucleic acid), which has a double-helix structure made of 4 molecules: adenine (A), guanine (G), cytosine (C), and thymine (T). Adenine on one strand of the double-helix always pairs with thymine on the complementary strand, and guanine always pairs with cytosine. 99.9% of our DNA is identical across humans, but on some sites of the DNA, variations can occur. These sites are called SNPs (single nucleotide polymorphisms). On these sites, some individuals may have an A-T combination, others may have a G-C combination, for example. To understand which of these variations are linked to which variation in outcomes, researchers use Genome-Wide Association studies (GWAS). These studies have uncovered the SNPs which are most relevant in predicting outcomes such as BMI, mental health, education, height etc. In order to predict an individual's risk of a certain outcome, researchers construct polygenic scores - a weighted linear combination of individual's SNPs at relevant sites, where the weights are established from the GWAS. In the following, we use the polygenic score for BMI as a measure of individuals' genetic predisposition for BMI. More details can be found in appendix B.

3.2 Descriptives

The relationship between BMI and genes Figure 1 shows the association between the genetic predisposition for BMI as measured by the BMI polygenic score and BMI at different ages. We can see that whilst at ages 3 and 5, the relationship is not particularly tight, the relationship becomes tighter over time. In particular, the R^2 rises from 0.33 at age 3 to 0.84 at age 11. Additionally, the gradient between BMI PGS and BMI becomes steeper. This is in line with evidence from the genetics literature that the influence of genes becomes stronger with age (Bouchard, 2013; Tucker-Drob et al., 2013). Houmark et al. (2024) also find that the predictive power of the educational attainment polygenic score for cognitive skills increases as children age.

The relationship between cohort member genes and parental genes Table 1 shows correlations between the cohort member's polygenic scores, their mother's polygenic score and their father's polygenic scores. We can see that maternal (paternal) BMI PGS increases the cohort member's BMI PGS by around 0.5, conditional on paternal (maternal) BMI. This makes sense, as in expectation, the genes of the child are an average of the mother's and the father's genes (see Appendix B for further elaboration on this). Moreover, we can see that the R^2 is also around 0.5. Because conditional on parental genes, the child's genes are random, half of the variation of a child's genes is explained by variation in parental genes, whereas the other half of the variation is the variation conditional on parental genes. Finally, the last row

Figure 1: BMI and BMI PGS



of Table 1 shows the correlation between maternal and paternal polygenic scores. We can see that the correlation is relatively low, indicating that there is little assortative matching based on BMI polygenic scores. Those with low (high) BMI PGS do not seem to particularly seek out partners that also have a low (high) BMI PGS.

Table 1: OLS regression of cohort member polygenic scores on maternal and paternal polygenic scores

	Cohort member: BMI PGS
Mum: BMI PGS	0.506***
	(0.014)
Dad: BMI PGS	0.518***
	(0.014)
R^2	0.520
Correlation between mum's PGS & dad's PGS	-0.021

Obesity and Genes The above description has primarily focused on BMI as an outcome. However, variations in BMI within a healthy range are rarely of concern. Instead, obesity is what health practitioners and policy-makers tend to be concerned about - given its significance costs, at both individual and society level. Turning to obesity, Table 2 shows that obesity rates increase with age. By the time the cohort reaches age 17, 10% of cohort members are obese. To gain an initial understanding as to whether BMI PGS matters for obesity, we split the BMI PGS into quartiles. Focusing on obese children, we then show the share of children belonging to each BMI PGS quartile. Figure 2 shows that those who are obese are

much more likely to belong to higher BMI PGS quartiles. For example, at age 3, 33% of obese children were in the highest BMI PGS quartile, whereas only 16% where in the lower BMI PGS quartile. This pattern becomes starker at age 17. At age 17, almost half of the children who are obese have a high BMI PGS. This indicates a strong relationship between BMI PGS and obesity.

Conversely, Figure 3 shows that of those in the highest quartile of BMI PGS, 20% are obese at age 17, whereas of those in the lowest quartile of BMI PGS, less than 5% are.

Table 2: Obesity rates across different ages

	Mean	SD
Age 3	0.058	0.233
Age 5	0.058	0.233
Age 7	0.061	0.239
Age 11	0.067	0.250
Age 14	0.074	0.263
Age 17	0.105	0.307

Investments and Genes Finally, we investigate the relationship between the BMI PGS and our investments – diet quality and exercise. To do so, we bin BMI PGS into vintiles and calculate the average diet quality, exercise, and BMI for each vintile (standardized and pooled across all ages). We also estimate univariate regressions of diet quality, exercise, and BMI on BMI PGS. The left-most panel of Figure 4 plots the average diet quality against vintiles of the BMI PGS and also reports the coefficient from the univariate regression (with standard errors in parentheses). The middle panel and right-most panel analogously show the associations between exercise and BMI PGS, and BMI and BMI PGS. We can see that there are significant and negative associations between BMI PGS and investments. A standard deviation increase in BMI PGS is associated with a 0.05 of a standard deviation decrease in diet, and a 0.03 of a standard deviation decrease in exercise. Individuals with higher BMI PGS receive fewer health investments during childhood. The magnitudes of these associations, however, are much smaller compared to, for example the association between standardized BMI and BMI PGS. A standard deviation increase in BMI PGS is associated with a 0.23 standard deviation increase in BMI.

To summarize, in this section we have established that 1) the influence of genes on BMI becomes stronger as children become older, 2) cohort member PGS is conditionally random given parental PGS, and there is no assortative mating based on PGS, 3) children who are obese are much more likely to have a high BMI PGS as compared to children who aren't obese, 4) higher BMI PGS is associated with worse diets and less exercise, though the associations are not very large in magnitude.

We will now study in more detail the production of BMI and obesity throughout childhood.

Figure 2: Quartile of BMI PGS conditional on being obese

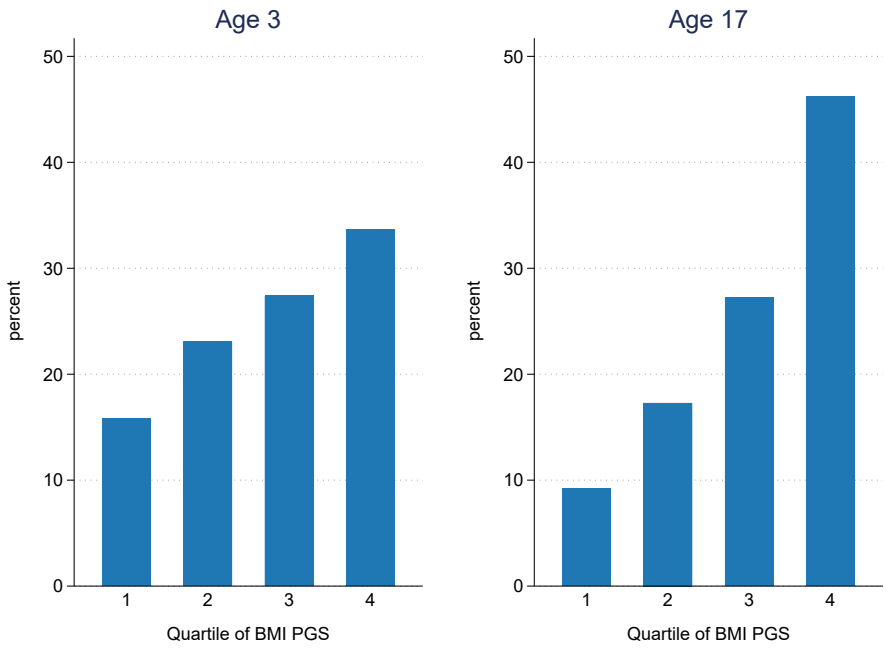


Figure 3: Obesity rates conditional on quartile of BMI PGS

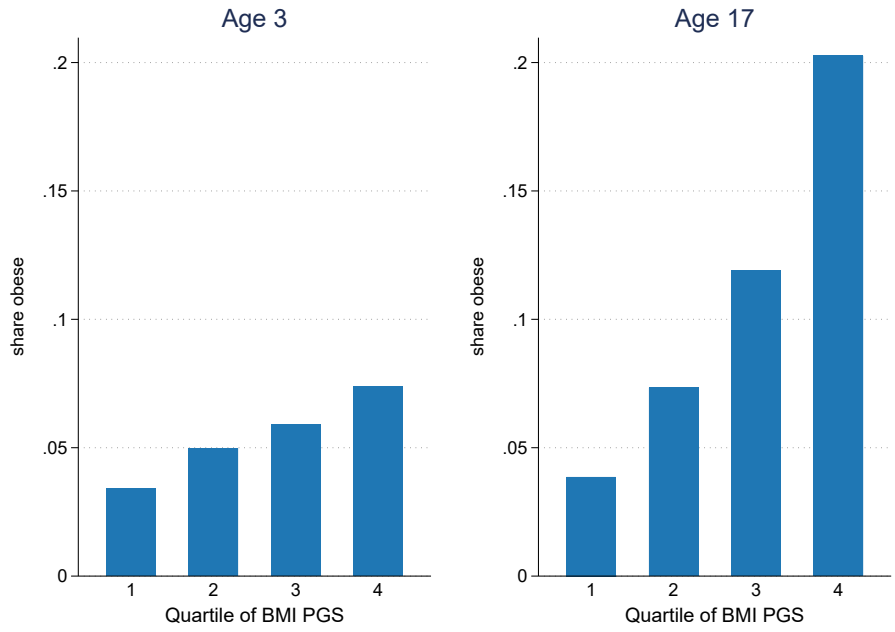
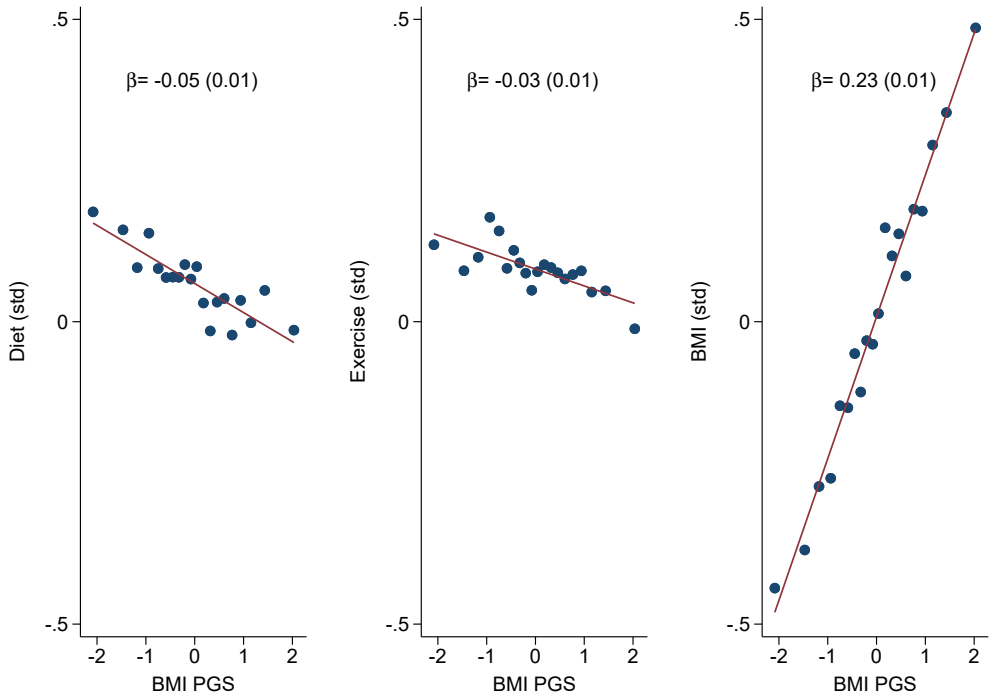


Figure 4: Associations between BMI PGS and standardized diet, exercise, and BMI.



4 Framework: The role of genes in BMI formation

To understand how genes can influence BMI and obesity, we use a simple framework consisting of production functions of BMI and obesity, as well as investment equations for diet and exercise.

Production functions We model the production function of BMI to be:

$$B_{it} = \alpha_{0,t} + \alpha_{1,t} B_{it-1} + \alpha_{2,t} d_{it-1} + \alpha_{3,t} ex_{it-1} + \alpha_{4,t} G_i + \boldsymbol{\alpha}_{5,t} \boldsymbol{B}^p + \boldsymbol{\alpha}_{6,t} \boldsymbol{G}^p + \boldsymbol{\alpha_{5,t}} \boldsymbol{X_{it}} + \epsilon_{it}^B$$

where BMI (B) is a function of past BMI, diet (d) and exercise (ex), BMI polygenic score (G_i), a vector of parental BMI ($\boldsymbol{B}^p = [B^m, B^f]$), a vector of parental BMI PGS ($\boldsymbol{G}^p = [G^m, G^f]$) and other controls $\boldsymbol{X_{it}}$. Several things should be noted: First, the model allows for a direct, biological effect of genes on BMI (α_4): Those with a high BMI PGS may have a differently built body or a different metabolism. Second, in theory, BMI is a function of only genes, past BMI, calories taken in via one's diet and calories out via one's exercise (Bouchard, 2021; Biroli, 2015; Yang et al., 2007). However, we do not exactly observe calories taken in, but instead observe a measure of quality of an individual's diet. Similarly, we do not exactly observe calories expedited via exercise, but instead we know how often the individual exercises. To account for the fact that our diet and exercise measures do not fully capture calories taken in or expedited,

we further control for parents' BMI and parents' BMI PGS which are likely strongly correlated with the food environment and exercise habits at home. Additionally, we control for the number of sibilings, region, and family income. These variables capture *family genetic associations*, that is a correlation between BMI PGS and BMI that is not causal, but instead arises because the families and environments of those with high BMI polygenic scores are different from those with low BMI polygenic scores.

The model is analogous for obesity, but uses a logit form due to the outcome being binary:

$$\ln\left(\frac{Pr(obese)}{1-Pr(obese)}\right) = \beta_{0,t} + \beta_{1,t}\mathbb{1}_{obese_{t-1}} + \beta_{2,t}d_{it-1} + \beta_{3,t}ex_{it-1} + \beta_{4,t}G_i + \boldsymbol{\beta}_{5,t}\boldsymbol{B}^p + \boldsymbol{\beta}_{6,t}\boldsymbol{G}^p + \boldsymbol{\beta_{5,t}X_{it}}$$

Investment equations The key investments we allow for are diet and exercise. We model diet and exercise using the following equations:

$$\begin{aligned} d_{it} =& \gamma^d_{0,t} + \gamma^d_{1,t}\ln B_{it} + \gamma^d_{2,t}G_i + \gamma^d_{3,t}\boldsymbol{B}^p + \gamma^d_{4,t}\boldsymbol{G}^p + \gamma^d_{5,t}P_{it} + \boldsymbol{\gamma^d_{6,t}X_{it}} + \epsilon^d_{it} \\ ex_{it} =& \gamma^e_{0,t} + \gamma^e_{1,t}\ln B_{it} + \gamma^e_{2,t}G_i + \gamma^e_{3,t}\boldsymbol{B}^p + \gamma^e_{4,t}\boldsymbol{G}^p + \gamma^e_{5,t}D_{it} + \boldsymbol{\gamma^e_{6,t}X_{it}} + \epsilon^d_{it} \end{aligned}$$

We model diet and exercise as functions of current BMI, BMI polygenic score, BMI of the parents, and our vector of controls (number of siblings, family income, region of residence). Additionally, we include two instruments that are excluded from the BMI production function: We allow diet to be affected by food prices (P) and exercise by the distance to parks and leisure facilities (D).

In addition to the *direct* biological effect of genes mentioned above, the investment equations allow for another *indirect* effect that Houmark et al. (2024) call *nurture of nature*: Parents might reinforce, or compensate for, children's genetic makeup by investing more or less in those with high BMI or high BMI PGS. An individual's genes may influence their diet and exercise, and thus BMI. This *nurture of nature* effect (the influence of genes on investments) is captured by parameters $\gamma^d_{1,t}, \gamma^d_{2,t}$ and $\gamma^e_{1,t}, \gamma^e_{2,t}$.

Again, family genetic associations come into play here. Those with a high BMI PGS may happen to live in regions with poor quality school meals or are more likely to have parents with a high BMI PGS who may provide a worse diet at home (captured by $\gamma^d_{3,t}$ and $\gamma^d_{4,t}$) or exercise less ($\gamma^e_{3,t}$ and $\gamma^e_{4,t}$).

5 Results

In the following, we first show some preliminary evidence as to how BMI PGS is associated with BMI, diet, and exercise. We then turn to the full production function estimates.

Preliminary Evidence Table 3 shows that the association between BMI PGS and BMI increases over time, in line with the previous literature. Whilst at age 7, a standard deviation increase in BMI PGS is

associated with a BMI increase of 0.5 points, at age 17, the association increases to 1.4 points.

Once we include parental BMI, we can see that this association decreases.[1] Both maternal and paternal BMI significantly predict BMI during all childhood ages. This indicates that whilst a large part of the association between BMI PGS and BMI comes from a direct effect of BMI PGS on BMI, an important share also stems from family genetic associations and the nurture of nature. We will present a detailed decomposition of the importance of these different channels in subsection 5.1. Importantly, compared to the cognitive skill formation process for which around half of the genetic association between genes and cognitive skills can be attributed to a direct genetic effect, Table 3 indicates that for the BMI formation process, a much larger fraction can be attributed to a direct genetic effect.

Turning to investments we can see significant negative associations between BMI PGS and diet and exercise for almost all ages. These are attenuated by around 1/2, once we include maternal and paternal BMI. Interestingly, for diet, at ages 7, 14, and 17, even if we find some attenuation, the effect of BMI PGS continues to be significant, indicating that there may be a direct link between individuals' genes and their diets. In contrast, we do not find such a link for exercise. Except for age 7, once we include parental BMI, the association between genes and exercise ceases to be significantly different from zero. This is in line with Biroli (2015), who also finds a direct link between individuals' diet and their genes, but not for exercise.

Table 3: BMI PGS and BMI

	BMI 7	BMI 11	BMI 14	BMI 17
Panel A				
BMI PGS	0.497***	0.947***	1.125***	1.402***
	(0.028)	(0.044)	(0.051)	(0.066)
N	5977	6032	6238	4956
r2	0.054	0.074	0.075	0.090
Panel B				
BMI PGS	0.373***	0.709***	0.842***	1.054***
	(0.027)	(0.042)	(0.048)	(0.063)
BMI Mom	0.098***	0.197***	0.237***	0.272***
	(0.007)	(0.011)	(0.012)	(0.015)
BMI Dad	0.096***	0.175***	0.208***	0.239***
	(0.008)	(0.013)	(0.015)	(0.019)
N	5,977	6,032	6238	4,956
r2	0.128	0.183	0.188	0.206

[1] We control for parental BMI here rather than parental PGS as there are fewer missings in parental BMI. In future work, we plan to impute missing parental PGS. Previous research has shown that accounting for parental phenotypes fully captures relevant variation in parental genotypes. For example, Houmark et al. (2024) find that controlling for parental education accounts for all relevant variation of parental EA PGS.

Table 4: BMI PGS and Diet

	Diet 7	Diet 11	Diet 14	Diet 17
BMI PGS	-0.050***	-0.031**	-0.052***	-0.071***
	(0.012)	(0.013)	(0.012)	(0.016)
N	6012	6120	6364	3714
r2	0.007	0.004	0.010	0.010
Panel C				
BMI PGS	-0.037***	-0.015	-0.032***	-0.047***
	(0.013)	(0.013)	(0.013)	(0.016)
BMI Mum	-0.011***	-0.013***	-0.017***	-0.024***
	(0.003)	(0.003)	(0.003)	(0.004)
BMI Dad	-0.002	-0.006*	-0.004	-0.002
	(0.004)	(0.004)	(0.004)	(0.005)
N	6,012	6,120	6,364	3,714
r2	0.023	0.023	0.035	0.037

Table 5: BMI PGS and Exercise

	Exerc 7	Exerc 11	Exerc 14	Exerc 17
BMI PGS	-0.055***	-0.055**	-0.050**	-0.033
	(0.018)	(0.023)	(0.026)	(0.032)
N	6017	6133	6386	5176
r2	0.003	0.005	0.002	0.002
Panel C				
BMI PGS	-0.036**	-0.027	-0.019	-0.007
	(0.018)	(0.023)	(0.027)	(0.032)
BMI Mum	-0.022***	-0.022***	-0.025***	-0.023***
	(0.004)	(0.005)	(0.006)	(0.007)
BMI Dad	0.010**	-0.009	-0.017**	-0.006
	(0.005)	(0.006)	(0.008)	(0.009)
N	6,017	6,133	6,386	5,176
r2	0.028	0.020	0.008	0.007

BMI production functions Table 6 shows the raw associations between BMI and BMI PGS, as well as the production functions of BMI in childhood for ages 7 to 17. The raw associations show what we have discussed earlier - the associations between genes and BMI become stronger over time. However, once we consider the production functions, we can see that the direct effect of genes is actually fairly constant over time. A standard deviation increase in BMI PGS directly increases BMI by between 0.214 to 0.287 points. One explanation of why a constant effect of BMI PGS may amplify over time is the high persistence of BMI over time that we can also see in Table 6, which ranges between 0.868 to 1.18. For example, a one standard deviation increase in BMI PGS leads to a 0.219 increase in BMI at age 7, and a 1.178*0.219=0.257 increase in BMI at age 11, solely due to the persistence of age 7 BMI. Turning to

parental genes and parental BMI, we observe that both maternal and paternal BMI have a statistically significant effect on BMI, even conditional on BMI PGS and parental BMI PGS. This result is similar to Houmark et al. (2024), who find that conditional on parental phenotype, parental genotype does not further predict the outcome of interest. The fact that BMI of the mother and father are important predictors of BMI, conditional on a child's genes indicates that the influence of parents on the BMI of their children goes beyond simply transmitting their genes. Finally, we turn to investments. We find that diet quality and exercise significantly reduce BMI at age 11. At other ages, diet quality also has a negative, but noisily estimated effect on BMI. The results for exercise are mixed, at age 17, we can find a positive and significant association of exercise on BMI. However, this may also be due to exercise building muscle-mass, which also is reflected in a higher BMI. Indeed, in Table 9 in the Appendix, we see that more frequent exercise has a negative and significant conditional association with the probability of being obese at ages 11, 14 and 17.

To quantify how much of the association between BMI PGS and BMI is explained by the direct effect, the nurture of nature, and by family genetic associations, we conduct a decomposition in section 5.1.

Results for obesity are qualitatively similar and can be found in Appendix Table 9.

Investment Equations Table 7 shows the estimates for our investment equations. For diet, we find a significant effect of BMI PGS even when we control for our full set of covariates, indicating that indeed, genes may directly affect dietary choices. The effect, however, is relatively small. For exercise, in contrast, whilst we can see a significant and negative association between BMI PGS and exercise at every age, there is no direct effect of BMI PGS on exercise in the full model. Genes predicting BMI do not affect exercise behavior.

Whilst the effect of genes on diet and exercise is thus rather small, family background variables matter a lot. In particular, higher income and higher parental education predict higher investments, both in terms of diet and in terms of exercise. Increasing parental income by 10% increases diet quality by 1.4% and 3.9% of a standard deviation. Similarly, a 10% increase in parental income increases diet quality by 2.5% and 3.8% of a standard deviation. In terms of parental education, children of mothers with a university degree have diet quality that is between 29% and 51% standard deviations higher than those who left school with GCSEs (age 16). Interestingly, both, the education of the mother and father predict investments.

Finally, parental BMI only matters for predicting diet at age 17 and exercise at ages 5 and 7. At older ages, the cohort member's own BMI starts to become more predictive of diet and exercise behavior.

Table 6: BMI production functions

	BMI 7	BMI 7	BMI 11	BMI 11	BMI 14	BMI 14	BMI 17	BMI 17
BMI PGS	0.502***	0.219***	0.957***	0.285***	1.127***	0.214***	1.374***	0.287***
	(0.029)	(0.029)	(0.045)	(0.036)	(0.052)	(0.038)	(0.065)	(0.046)
BMI		0.868***		1.178***		0.942***		0.940***
		(0.050)		(0.024)		(0.014)		(0.013)
Diet		-0.026		-0.061**		-0.033		-0.003
		(0.024)		(0.031)		(0.030)		(0.038)
Exercise		0.025		-0.053***		0.018		0.041**
		(0.018)		(0.020)		(0.015)		(0.018)
Mum: BMI PGS		-0.032		-0.057		-0.024		-0.001
		(0.025)		(0.037)		(0.039)		(0.046)
Dad: BMI PGS		-0.058**		0.023		-0.057		-0.021
		(0.028)		(0.044)		(0.044)		(0.052)
BMI Mum		0.044***		0.081***		0.053***		0.051***
		(0.006)		(0.007)		(0.007)		(0.009)
BMI Dad		0.043***		0.067***		0.053***		0.031***
		(0.007)		(0.009)		(0.009)		(0.012)
Controls								
N	5,758	5,758	5,572	5,572	5,865	5,865	4,842	4,842

Table 7: Investment equation

	Diet 5	Diet 5	Diet 7	Diet 7	Diet 11	Diet 11	Diet 14	Diet 14
Panel A								
BMI PGS	-0.045***	-0.014	-0.053***	-0.023**	-0.030**	0.014	-0.056***	-0.022*
	(0.012)	(0.012)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)
BMI		0.008		0.001		-0.009**		0.010***
		(0.008)		(0.006)		(0.004)		(0.003)
BMI Mum		-0.001		-0.003		-0.002		-0.011***
		(0.003)		(0.003)		(0.003)		(0.003)
BMI Dad		-0.003		-0.002		-0.005		-0.009**
		(0.004)		(0.004)		(0.004)		(0.004)
Income (ln)		0.135***		0.157***		0.282***		0.392***
		(0.024)		(0.026)		(0.041)		(0.040)
High School		0.301***		0.283***		0.200***		0.160***
		(0.036)		(0.037)		(0.038)		(0.038)
College		0.506***		0.469***		0.382***		0.294***
		(0.041)		(0.042)		(0.043)		(0.043)
High School		0.132***		0.107***		0.100***		0.169***
		(0.035)		(0.035)		(0.035)		(0.036)
College		0.208***		0.190***		0.230***		0.396***
		(0.039)		(0.038)		(0.039)		(0.040)
	Exercise 5	Exercise 5	Exercise 7	Exercise 7	Exercise 11	Exercise 11	Exercise 14	Exercise 14
Panel B								
BMI PGS	-0.028**	0.006	-0.039***	-0.002	-0.032**	0.011	-0.027**	0.011
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
BMI		0.009		-0.002		-0.011***		-0.027***
		(0.008)		(0.006)		(0.004)		(0.003)
BMI Mum		-0.008***		-0.005**		-0.003		-0.004
		(0.003)		(0.003)		(0.003)		(0.003)
BMI Dad		0.008**		0.007**		-0.004		-0.001
		(0.004)		(0.004)		(0.004)		(0.004)
Income (ln)		0.252***		0.338***		0.380***		0.047
		(0.025)		(0.027)		(0.040)		(0.044)
High school		0.195***		0.193***		0.109***		0.056
		(0.034)		(0.033)		(0.035)		(0.043)
College		0.319***		0.375***		0.222***		0.091**
		(0.042)		(0.041)		(0.041)		(0.047)
High school		0.112***		0.124***		0.117***		0.011
		(0.034)		(0.035)		(0.035)		(0.038)
College		0.282***		0.203***		0.211***		0.053
		(0.041)		(0.041)		(0.040)		(0.042)

5.1 Mediation analysis

As outlined above, there are three key mechanisms through which genetic endowments and BMI can be correlated: First, BMI PGS can affect BMI through a *direct*, biological mechanism. Second, there can be *nurture of nature*, i.e. those with a high BMI polygenic score have worse diets and exercise less. In childhood, this may be because they prod their parents to feed them more sweets or junk food, or refuse to participate in exercise. Lastly, there can be *family genetic associations*, which means that those with a high BMI PGS may live with obese parents who provide a worse food environment or attend schools with poor quality school meals. To estimate how much each of these three channels contributes to the association between BMI and BMI PGS that we observe, we conduct a Gelbach decomposition and show the results in Table 8. Details of the decomposition method can be found in appendix B.

Table 8 makes evident that the main driver of the association between BMI and BMI PGS are direct effects, which account for almost 3/4 of the association: those with higher BMI PGS have a higher BMI because of the biological impact of their genes. This is in stark contrast to cognitive skills, for which Houmark et al. (2024) have found direct effects to account for at most 56% of the gene-skill association.

Nurture of nature does not account for any of the link between BMI PGS and BMI. This is unsurprising given the weak relationships between BMI PGS and diet and exercise we found. It is important to note here that polygenic scores are a limited indicator of genetic endowments, and that as the predictive power of polygenic scores improves, stronger links between investments and BMI PGS might be found in the future.

Lastly, family genetic associations account for around 27% of the association between BMI PGS and BMI at all ages. Thus, a bit more than a quarter of the association between genes and BMI is due to those with a certain genetic makeup living under circumstances that are conducive to weight gain, such as high parental BMIs, poverty, and regional differences.

Table 8: Mechanism Decomposition: Association between BMI PGS and BMI

	Age 7	Age 11	Age 14	Age 17
Direct effect	73.7%	72.8%	73.1%	74.3%
Nurture of Nature	-0.2%	0.5%	0.3%	-0.3%
Family genetic associations	26.6%	26.7%	26.6%	26.0%

6 Conclusion

In this paper, we have estimated the first BMI production function which incorporates *both* child and parental genetic and environmental inputs into a dynamic framework. We have documented that genetic

predisposition exerts a cumulative influence on BMI as children grow older – reinforcing findings from genetics literature and extending them into longitudinal childhood development.

We have also found that parental BMI matter both economically and statistically for child BMI, even after controlling for the child's genetic endowment. Moreover, we have observed important differences in how child behaviors and lifestyles are affected by parents' BMI. While a mother's BMI influences both a child's dietary behaviors and exercise activities, a father's BMI primarily affects a child's exercise activities.

Our results quantify the effects of genes on child BMI during childhood development, and also show that parents' BMI matter for child BMI and related behaviors. Given that parental BMI influences child BMI beyond genetic transmission, targeted interventions focusing on family-based dietary and exercise habits (starting in the early years) may offer effective strategies for addressing childhood obesity.

References

Amin, V., C. A. Flores, and A. Flores-Lagunes (2020, August). The impact of BMI on mental health: Further evidence from genetic markers. *Economics & Human Biology 38*, 100895.

Bairoliya, N., R. Miller, and V. M. Nygaard (2024, February). Exercise or Extra Fries? How Behaviors Impact Health Over the Life Cycle.

Biroli, P. (2015). Genetic and Economic Interaction in the Formation of Human Capital: The Case of Obesity.

Bouchard, C. (2021, May). Genetics of Obesity: What We Have Learned Over Decades of Research. *Obesity (Silver Spring, Md.) 29*(5), 802–820.

Bouchard, T. J. (2013, October). The Wilson Effect: the increase in heritability of IQ with age. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies 16*(5), 923–930.

Brittain, E. L., L. Han, J. Annis, H. Master, A. Hughes, D. M. Roden, P. A. Harris, and D. M. Ruderfer (2024, March). Physical Activity and Incident Obesity Across the Spectrum of Genetic Risk for Obesity. *JAMA Network Open 7*(3), e243821.

Cawley, J., E. Han, J. Kim, and E. C. Norton (2019). Testing for family influences on obesity: The role of genetic nurture. *Health Economics 28*(7), 937–952. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.3889.

Cawley, J., E. Han, and E. C. Norton (2011, August). The validity of genes related to neurotransmitters as instrumental variables. *Health Economics 20*(8), 884–888.

Fan, M., Y. Jin, and M. Zhang (2023, June). Genetic risk, childhood obesity, and educational achievements. *Economics of Education Review 94*, 102408.

Fang, M. Z. (2014). Violating the monotonicity condition for instrumental variable-dimorphic patterns of gene-behavior association. *Economics Letters 122*(1), 59–63.

Houmark, M. A., V. Ronda, and M. Rosholm (2024). The nurture of nature and the nature of nurture: How genes and investments interact in the formation of skills. *American Economic Review 114*(2), 385–425.

Libuy, N., D. Bann, and E. Fitzsimons (2021, November). Inequalities in body mass index, diet and physical activity in the UK: Longitudinal evidence across childhood and adolescence. *SSM - Population Health 16*, 100978.

Muslimova, D., H. van Kippersluis, C. A. Rietveld, S. von Hinke, and F. Meddens (2020, December). Dynamic Complementarity in Skill Production: Evidence From Genetic Endowments and Birth Order.

National Health System (2024). Children's overweight and obesity. NHS England Digital.

Silventoinen, K. and H. Konttinen (2020, February). Obesity and eating behavior from the perspective of twin and genetic research. *Neuroscience & Biobehavioral Reviews 109*, 150–165.

Tucker-Drob, E. M., D. A. Briley, and K. P. Harden (2013, October). Genetic and Environmental Influences on Cognition Across Development and Context. *Current Directions in Psychological Science 22*(5), 349–355.

VanderWeele, T. J., E. J. Tchetgen Tchetgen, M. Cornelis, and P. Kraft (2014, May). Methodological challenges in Mendelian randomization. *Epidemiology (Cambridge, Mass.) 25*(3), 427–435.

Willage, B. (2018, January). The effect of weight on mental health: New evidence using genetic IVs. *Journal of Health Economics 57*, 113–130.

World Health Organization (2024). Obesity and overweight. [fact sheet].

Yang, W., T. Kelly, and J. He (2007, January). Genetic Epidemiology of Obesity. *Epidemiologic Reviews 29*(1), 49–61.

A Further Results

Table 9: Obesity production functions: Logit estimates

	Obese 7	Obese 7	Obese 11	Obese 11	Obese 14	Obese 14	Obese 17	Obese 17
F.obese								
BMI PGS	0.552***	0.325***	0.637***	0.387***	0.679***	0.468***	0.696***	0.440***
	(0.058)	(0.105)	(0.055)	(0.093)	(0.052)	(0.081)	(0.051)	(0.078)
obese		4.440***		3.915***		3.738***		3.589***
		(0.176)		(0.171)		(0.158)		(0.167)
diet		0.006		-0.113		-0.096		-0.039
		(0.089)		(0.075)		(0.069)		(0.064)
exercise_c		-0.088		-0.109*		-0.093**		-0.098***
		(0.067)		(0.061)		(0.039)		(0.032)
Mum: BMI PGS		0.053		0.025		0.017		0.010
		(0.108)		(0.092)		(0.085)		(0.079)
Dad: BMI PGS		-0.204		0.046		-0.035		0.189**
		(0.128)		(0.125)		(0.102)		(0.100)
bmimum		0.074***		0.092***		0.091***		0.077***
		(0.016)		(0.013)		(0.012)		(0.014)
bmidad		0.095***		0.097***		0.070***		0.085***
		(0.020)		(0.018)		(0.017)		(0.018)
Controls								
N	5758	5747	5571	5558	5865	5851	4842	4833

Table 10: BMI PGS and Exercise: non-linear

	Exerc 7	Exerc 11	Exerc 14	Exerc 17
4 quantiles of bmipgs_CM=2	-0.031	-0.122**	-0.067	-0.106
	(0.052)	(0.065)	(0.073)	(0.089)
4 quantiles of bmipgs_CM=3	-0.080	-0.094	-0.092	-0.067
	(0.052)	(0.066)	(0.073)	(0.088)
4 quantiles of bmipgs_CM=4	-0.141***	-0.169**	-0.105	-0.089
	(0.051)	(0.065)	(0.074)	(0.089)
N	6017	6133	6386	5176
r2	0.003	0.006	0.002	0.002
Panel C				
4 quantiles of bmipgs_CM=2	-0.014	-0.098	-0.042	-0.083
	(0.052)	(0.065)	(0.073)	(0.089)
4 quantiles of bmipgs_CM=3	-0.040	-0.036	-0.038	-0.021
	(0.052)	(0.066)	(0.073)	(0.089)
4 quantiles of bmipgs_CM=4	-0.093*	-0.098	-0.030	-0.020
	(0.052)	(0.066)	(0.075)	(0.090)
bmimum	-0.022***	-0.022***	-0.025***	-0.023***
	(0.004)	(0.005)	(0.006)	(0.007)
bmidad	0.010**	-0.010	-0.018**	-0.007
	(0.005)	(0.006)	(0.008)	(0.009)
N	6017	6133	6386	5176
r2	0.028	0.020	0.008	0.007

Table 11: BMI PGS and diet: non-linear

	Diet 7	Diet 11	Diet 14	Diet 17
4 quantiles of bmipgs_CM=2	-0.045	-0.028	-0.081**	-0.081*
	(0.035)	(0.036)	(0.036)	(0.045)
4 quantiles of bmipgs_CM=3	-0.122***	-0.098***	-0.103***	-0.102**
	(0.036)	(0.035)	(0.036)	(0.048)
4 quantiles of bmipgs_CM=4	-0.118***	-0.081**	-0.138***	-0.186***
	(0.035)	(0.036)	(0.036)	(0.046)
N	6012	6120	6364	3714
r2	0.008	0.005	0.009	0.010
Panel C				
4 quantiles of bmipgs_CM=2	-0.033	-0.016	-0.065**	-0.063
	(0.035)	(0.035)	(0.035)	(0.045)
4 quantiles of bmipgs_CM=3	-0.094***	-0.065**	-0.064*	-0.056
	(0.036)	(0.035)	(0.036)	(0.047)
4 quantiles of bmipgs_CM=4	-0.084**	-0.040	-0.088**	-0.123***
	(0.036)	(0.036)	(0.036)	(0.046)
bmimum	-0.011***	-0.013***	-0.018***	-0.024***
	(0.003)	(0.003)	(0.003)	(0.004)
bmidad	-0.002	-0.006	-0.004	-0.002
	(0.004)	(0.004)	(0.004)	(0.005)
N	6012	6120	6364	3714
r2	0.024	0.023	0.035	0.037

B Further Details