Intergenerational transmission of human body size: the role of genetics and culture

Alberto Palloni[1,2], Néstor Aldea-Ramos[1,3], and Hiram Beltrán-Sánchez[4]

[1]ECHO project, Spanish Research Council, Madrid, Spain
[2]Center for Demography and Ecology, University of Wisconsin-Madison, Madison, WI, USA
[3]French Institute for Demographic Studies (INED), Aubervilliers, France
[4]Fielding School of Public Health and California Center for Population Research, UCLA, Los Angeles, CA, USA

June 30, 2025

Shortitle: Models of human obesity

Keywords: Obesity, Polygenic Risk Score, Cultural Risk Score, Assortative mating, Differential Fitness,

*This is ongoing and unpublished work. Please do not cite or share without permission.

Abstract

As is the case of most human phenotypes, the trajectory of body size is a function of genetic transmission, gene environment interactions, assortative mating, and net fertility differentials. In this paper we demonstrate that, as predicted in recent population genetics models on niche construction, two additional forms of inheritance, vertical cultural (from parents to offspring) and horizontal cultural (from peers) are needed to account for evolution of human body size. To partially capture these transmission pathways, we introduce the notion of cultural risk score (CRS) that refers to the effects of a parental household niche on children's body size. We use a microsimulation to quantify the relative effects of each transmission mechanism and demonstrate the dominant role played by the CRS.

1 Introduction

There is a rich tradition in population genetics that establishes the foundations of models for the diffusion of a phenotype in human populations (1; 2; 3; 4; 5; 6). The nature of these models is heterogeneous, reflecting differences in the problems they address. A majority of these models are designed to assess influences of selection, drift, mutation, and migration on the time and space distribution of a trait. They introduce consideration of substrate processes that are the raw materials on which selection, drift, mutation, and migration work: genotypic and phenotypic transmission of the phenotype, assortative mating, and differential fitness (net fertility).

After 1980 or so, standard genetic models were significantly modified. Beginning with the seminal contribution by Cavalli-Sforza and Feldman (7), an important determinant of the evolution of human traits, cultural transmission, began to be included in most population models alongside genetic transmission. Recent models consider several substrate processes related to cultural inheritance of traits, from vertical (parent to offspring), to horizontal (from peers) to oblique (from teachers, mentors, and influencers). As expected, the complexity of the models increases in tandem with the number of substrate processes included. To circumvent reduced mathematical tractability without ignoring relevant processes, simplifying assumptions are introduced. Thus, for example, in most cases, the trait is binary, not continuous, genetic transmission is a function of a single or a handful of biallelic loci (6; 8; 9; 10; 11; 12), mating is for the most part assumed to be completely random or completely endogamous by phenotype, seldom by social groups or a combination thereof and, finally, differential fitness (net reproduction) is usually defined as a simple function of the phenotype, invariant over time and space. With some exceptions, these models assume that there are no gene-environment, GxE, interactions.

In this paper, we introduce a model for the diffusion of obesity that does not invoke many of the restrictive assumptions required in recent literature. We model a continuous phenotype (body mass index (BMI)), use continuous genetic values (Polygenic Risk Scores or PRS), allow for GxE interactions and consider mating by social class. Unlike most recent models of human traits, we include vertical and horizontal cultural transmission using an empirically based construct, the cultural risk score, CRS, reflecting the strength of non-genetic effects on children's body size

that emerge from environments shared by parents and offspring. We show that vertical cultural transmission, from parents to offspring, and horizontal cultural transmission (from peers), might play an important role in the diffusion of obesity. We demonstrate that cultural transmission can be a dominant force and that, in combination with fertility differentials and assortative mating, is likely to account for global observed time trends of obesity.

2 Nature of model

Below, we describe the domains included in the microsimulation model.

2.1 Genetic transmission(GT)

There are two strategies to account for GT. The first is to invoke Mendelian segregation rules and assume that the phenotype is dichotomous and fully determined by one or a few of biallelic loci. This strategy is appropriate for monogenic phenotypes but unrealistic for most phenotypes of interest to social and health scientists, such as Body Mass Index (BMI), Type 2 Diabetes (T2D), height, education, IQ, etc. We have shown elsewhere (13) that, at least in the case of obesity, a simplified biallelic model can lead to misleading inferences.

A better strategy is to exploit current advances in genetic computing and meta-analyses of massive genetic data bases (GWAS). These studies have powered the construction of PRSs that capture effects of hundreds of loci and render unnecessary simplifying assumptions about genotypes. We will use this approach and focus on a single phenotype, Body Mass Index (BMI) that will be defined as

$$BMI_i = \alpha + \beta \times PRS_i + \gamma \cdot CRS_i + \epsilon_{i,BMI} \tag{2.1}$$

where BMI_i, PRS_i and CRS_i are the BMI, PRS and CRS of an individual i, α, β and γ are parameters to be estimated and $\epsilon_{i,BMI}$ is a $\mathcal{N}(0, \sigma_{BMI})$ normal variate.

2.2 Gene-Environment interactions (GxE)

A phenotype trajectory can shift when genetic effects on a phenotype change more than trivially with environmental conditions. Examples of phenotypes associated with GxE interactions in social

and health sciences are body size (BMI and obesity), intelligence (IQ), and educational attainment. There has been a large body of recent research on GxE involving BMI and obesity. First, GxE interaction effects have been found in populations that experience sharp contrasts between their ancestral and modern environments. US native populations experience unexpectedly high loads of obesity given their genetic profile (14). Existence of mismatches is also an explanation for the excess obesity detected in migrant populations from low income to high income countries despite similarity in their genetic makeup (15). But even in settings where chances of mismatches are small, environments shaped by external influences may strengthen (attenuate) the impact of genetic propensities. Thus, for example, when placed in modern obesogenic settings, entire US and UK birth cohorts show larger than expected increases in BMI given their individual genotypes (16; 17). Note that the overall impact of GxE interactions, however, has a limited reach. They certainly can influence aggregate short-term trends in BMI and obesity but, in the absence of strong selection, they cannot change allelic frequencies. Depending on the AM regime, they could at most alter the genotypic composition by phenotype of subsequent generations. To accommodate GxE interactions we use the following definition

$$GxE = \begin{cases} I \times PRS \times CRS, & \text{if PRS>0, CRS>0} \\ 0, & \text{otherwise} \end{cases} \tag{2.2}$$

where I is the magnitude of the interaction, taking values between 0 and 2[1]. This term is added to equation (2.1) and allows to account for GxE and also reproduces the right-hand thick tail of the BMI distribution observed in Western populations (18).

2.3 Vertical cultural transmission (VCT)

Although VCT is of importance in some animal species other than humans, it is uniquely influential for the inheritance of human phenotypes of interest. With few exceptions (7; 5), the classic population literature, eschewed this issue altogether. However, more recently, a number of researchers, particularly Feldman and colleagues (8; 9; 10; 11; 12), have made significant contributions in this

[1] A number of very recent studies shows that increased penetrance due to GxE can be as low as .02 and as high as .4 (17)

area.

We follow in their steps and implement a strategy that includes the influence of culture on body size in a way analogous to the PRS (19). In particular, we estimate Structural Equation Models (SEM) for the BMI of children who live in households and about whom we possess full phenotypical and genotypical (PRSs) information about parents-children pairs as well as of domains such as household diet, physical activities, sleep, and stress. One of the ancillary products of SEMs is expression of a latent variable that reflects the effects of households' setting, parental and children PRS, and indirect genetic effects (IGE) on children's body size. This is the cultural risk score (CRS), a normalized quantity that can be used in exactly the same way as a PRS is used in the expression for BMI. Instead of reflecting genetic propensities to body size, the CRS reflects household, niche-related propensities. Our model accounts for this using expression (2.1). We assign alternative values to the γ parameter reflecting the strength of the effects on body size.

2.4 Horizontal cultural transmission (HCT)

Many studies have demonstrated the influence that peers can have on an individuals' body size (20; 21; 22; 23). In principle, it should be possible to translate peers' effects estimated in these literature into a metric of HCT analogous to the CRS. However, to do so, we would need to use the original data sets and estimate SEMs analogous to those implemented in the case of CRS. This require using the original data sets to which we have no access. To circumvent this obstacle and as a temporal solution, we proceed as follows: we assume that HCT operates by modifying the CRS inherited from parents at a crucial time in a generation's life, namely, right before they reproduce. We will also assume that the CRS that a parent generation transmits to their offspring is the one already modified by HCT: a weighted average of the CRS in their social class at maturity (at age 25-29, after social class mobility experiences) and of the one inherited from their parental household. The magnitude of the weight reflects the strength of HCT.

2.5 Assortative mating (AM)

The nature of the population's mating regime is crucial for the evolution of most human traits. Mating, or reproductive pair formation, can occur by phenotype, social homogamy (e.g. population

stratification), or a combination of these. The difference between these mechanisms is consequential for the propagation of the trait.

An AM regime driven by phenotype modulates the impact of GT. To the extent that couples that resemble each other phenotypically are more likely to be genetically similar (24; 25; 26), AM will influence both their offspring's genotype and phenotype distributions. When the phenotype is influenced by only a few loci with strong penetrance, the impact of AM on the trait's evolution could be strong. Instead, when genetic value is assessed with a PRS, the linkage between AM and genetic heredity is relaxed[2]. When mating is random relative to the phenotype, genotypic frequencies will remain constant across generations, and the trait's evolution will be a function of other forces (13)

AM does not only, or mainly, take place by phenotype, but also according to social homogamy rules. AM by social homogamy can modulate *vertical cultural* transmission (VCT) of the phenotype because membership in social groups results in shared environments and behaviors that might influence the phenotype (7; 5). In our case, the pathway that enables these influences to operate is via the CRS. Furthermore, social homogamy creates conditions under which phenotypic distributions are influenced by indirect genetic effects (IGE)[3]. As we show later, exposures to early familial environments contributes significantly to children's body size and these, in turn, are consequential for adult manifestations of the phenotype. Although there is empirical evidence showing that preferential mating by body size is quite common, social homogamy by social class might create the appearance of mating by phenotype [4]. When, as happens with BMI, the AM regime is a blend of both phenotypical and social homogamy rules, there could be feedback effects that strengthen the influence that GT has on the phenotype.

A highly controversial issue is whether changes in the AM regime can offset the impact of net fertility differentials. For example, it is known that there are strong net fertility differentials by intellectual ability (IQ), education, and body size (29; 30; 31; 32; 26). These differentials need not

[2]AM can also increase gametic linkage disequilibrium (GLD), the association between distant loci (27; 28) and thus augment the genetic variance of the trait.

[3]IGE refer to impacts on a focus individual (offspring) phenotype that is influenced by an environmental phenotype (parental household) which is, in turn, a function of parents or other kin genotypes.

[4]This could be a result of the fact that social class is strongly correlated (both negatively and positively) with body size.

be associated with the genotype or phenotype but could be a result of shared social contexts. Under what conditions can shifts in mating regimes (phenotypical or social homogamy) significantly alter the phenotype trajectories implied by regimes of differential net fertility? Below, we address this question in the case of BMI.

To represent assortative mating, there are three strategies. The first, assumes that mating is driven only by social class preferences and that any appearance of assortative mating by phenotype is the result of social homogamy by social class and a strong correlation between body size and social class. When there are only two social classes (as we will be the case in our model) we define a one-parameter continuous function, $\omega \in [0, 1]$ and assume all members of the population find a partner. When $\omega = 0$, the mating regime is completely random and when $\omega = 1$, the mating regime is completely endogamous. Other ω values within the closed interval [0,1] define mixed regimes.

Let p_{ii} be the fraction of pairs formed by individuals belonging to class i and π_i the fraction of individuals in subpopulation i. Then,

$$p_{ii} = \pi_i^2 \cdot (1 - \omega) + \pi_i \cdot \omega \tag{2.3}$$

for $i = 1, 2$ and

$$p_{ij} = \pi_i \cdot \pi_j \cdot (1 - \omega) \tag{2.4}$$

for $i \neq j$.

The second strategy assumes that mating is by phenotype only. If the phenotype is dichotomized, one can use expressions similar to those above. If more that two categories are needed, we simply increase the number of expressions of class (2.3) and (2.4) to capture k groups and as many as $k \cdot (k - 1)/2$ classes of distinct relevant pairs.

Finally, a third strategy is to model a blended form of mating regime. To implement this, populations are first paired by social class and then by phenotype. One can create subpopulations of pairs $(1, 1)$, $(2, 2)$ and $(1, 2)$. If the phenotype is also dichotomized, e.g. obese vs non-obese,

expressions (2.3) and (2.4) are applied with possibly different endogamy coefficients $\omega's$.[5].

In this paper we mainly discuss results from a model with AM regulated by social class membership. Even so, we test the effect of the nature of the AM regime (social class vs. phenotype)[6]. We hypothesize that, under some conditions to be specified later, when AM is by social class, the influence of VCT is more likely to have a strong impact than when it AM is by phenotype or a combination of both.

2.6 Net fertility differentials (FD)

We adopt the standard demographic definition of net reproduction rate of a reproductive couple c, NRR_c, assume that individual members of the pair are of the same age, and define the quantity as follows:

$$NRR_c = \int_a^b g_c(x)\kappa_c(x)dx \tag{2.5}$$

where $g_c(x)$ is the fertility rate of pair c at exact age x [7], $\kappa_c(x)$ is the couple's members joint probability of surviving to age x and a and b are the initial and final ages of reproduction. We simplify expression (2.5) to be $f_c = INT_c \cdot \kappa(A)$ where INT_c is the integral of the function $g_c(x)$ and A is mean age at childbearing[8].

Fertility differentials in our model will be associated with social classes, not with the phenotype. Since we work with dichotomized social class, the following expression will be used:

$$f_{ij}(t) = F(t) \cdot (1 + \varphi \cdot (1 - i - j)) \tag{2.6}$$

where $F(t)$ is the total fertility rate in time t, i and j are parental social classes (1 being upper class, and 0 lower class), and φ a one-parameter continuous function, $\varphi \in [0, 1]$ regulating fertility differentials between classes[9].

Importantly, fertility differentials will not only influence the phenotype distribution but also

[5] If there are more than two social classes and/or phenotypes, simple extensions of the above expressions apply.

[6] Extensions to cases in which AM is by phenotype are described elsewhere (13).

[7] Throughout, we assume that members of a pair are of the same age.

[8] Concentrating all fertility at the mean age of childbearing, does not bias results of a stable population model; it only limits the features of a phenotypical trajectory that can be studied.

[9] Note that expression (2.6) implies that $\varphi = 1$ maximizes fertility differentials while $\varphi = 0$ results in no differentials.

its time trend. From first demographic principles, we know that even if the net reproduction rate of a subpopulation with a given phenotype is reduced to 1 (implying a rate of growth equal to 0), the rate of growth of the subpopulation with that phenotype will continue to increase for some time afterward (33). Among other things, this implies that the distribution of the phenotype generated by any regime of net fertility differentials has a momentum of its own and that even if an intervention succeeds in changing fertility levels, its impact will not be felt until after some time has elapsed. The model we propose is equipped to assess the growth momentum of subpopulations whose NRR converges to 1.

3 Model description

We describe key features of the microsimulation model by following two generations, a parental one (G0) and their offspring (G1).

3.1 Parental generation, G0

We begin with generation G0, an arbitrary population of size Z individuals of which $Z/2$ are females with an age structure that replicates the one observed for the United States in 1950. Individuals are endowed with random values of PRS and a CRS, both drawn from two standard normal distributions implying that, at least at the outset, the PRS and CRS are uncorrelated. Each individual is then assigned to social classes (high/low) according to their CRS [10]. Finally, we use the following expression to define the parental generation BMI:

$$BMI_{i,x} = \alpha(x) + \beta(x) \times PRS_i + \gamma(x,c) \cdot CRS_i + GxE + \epsilon_{i,BMI} \tag{3.1}$$

where $\alpha(x)$ and $\beta(x)$ depend on age x, $\gamma(x,c)$ on age and social class c, and GxE is defined in (2.2).

3.2 Offspring generation, G1

Each pair in generation G0 has all their lifetime offspring at two ages, 25-29 and 30-34, half at each. The number of offspring for each pair is a function of their social class, as defined by (2.6).

[10] We use the relation between social class of parents and CRS estimated from the data sets used in (19)

These offspring become the members of generation G1. They are assigned a PRS equal to the average of their parents plus a normal random variate $\mathcal{N}(0, \sqrt{0}.5)$. Similarly, the offspring's CRS is the average of their parents plus a random variate $\mathcal{N}(0, (1-\nu) \cdot 2 \cdot \sqrt{0}.5)$, where $0 \leq \nu \leq 1$ stands for the fidelity of cultural parental transmission. When $\nu = 1$, offspring inherits exactly the parental average of the CRS. When $\nu < 1$, the transmission is distorted by randomness[11].

3.3 Social class mobility

The G1 adults' social class is not determined solely by their parents', but is affected by social class mobility. We consider a baseline mobility pattern based on the 1958 National Child Development Study in the UK. To simplify, the social class of destination is assigned at birth. Although this simplification implies that social class will have some effects during childhood *via* $\gamma(x, c)$ (see 3.1), these will be very small as their impact will be almost entirely manifested at adult ages (from 20-24, see below).

3.4 The life course of generation G1

Members of generation G1 are assumed to survive to age 25-29 with probability 1. They experience two changes. The first is that their BMI is updated, *via* the intercept α and by the genetic (β) and cultural (γ) penetrance (see expression (3.1)). α values are drawn from the 1958 National Child Development Study in the UK whereas β values are drawn by fitting a quadratic curve to the BMI-PRS to observed at multiple ages in that same cohort. γ values follow the same age pattern as β, but with higher values reflecting the magnitude of the effects of PRS and CRS on the BMI we obtained in previous research (19).

The second change concerns the influence of peers on their CRS. To account for HCT, we consider that at ages 20-24 individuals may be influenced by a class-specific (peer-specific) CRS derived from their social class of destination. As noted before, the levels of the association between CRS and social class are derived from previous work (19). The final adult CRS, that is, the one they will pass on to their offspring, is a weighted average between the inherited (VCT) and the class-specific one (HCT):

[11] We assume no bias in the transmission from parents to offspring

$$CRS_{adult} = a \cdot CRS_{child} + (1 - a) \cdot CRS_{class} \tag{3.2}$$

where a is a measure of the relative strength of parental and peer-influences CRS.

Finally, as their parents did before them, members of G1 form pairs according to social class homogamy rules (see above) and then procreate producing all their lifetime offspring associated with their social class.

3.5 Generation Gn

The simulation employs the information on class mobility, mating and reproduction of generation G1 and applies the algorithm described above using their offspring, namely, generation G2. The algorithm is then repeated for 50 time steps (five-year jumps). In each of those, a new generation is born and the population gets 5 years older.

4 Results

We address the following questions:

1. How strong is the impact of VCT compared to the other substrate processes? How different are outcomes in the case when there is no VCT?

2. How sensitive is a population's obesity prevalence to changes in AM and DF? Does AM have different impacts depending on whether or not there is (there is not) VCT? How large are these differences?

3. How influential is the prevailing class of AM regime? Do AM by phenotype and by social class induce important differences in trajectories of obesity prevalence?

4. How large can the impact of GxE interactions be on future trends of the phenotype?

In what follows, we answer these questions in turn.

4.1 The impact of VCT and ν

The left and right panels of Figure 1 are plots of obesity prevalence when VCT is ignored (left panel) and when it is fully included (right panel). When there is no VCT, differences in fidelity

(ν) have small effects, as it should be. However, when VCT is included, the contrasts in prevalence when ν is set at a minimum and a maximum are quite large. In fact, the maximum difference at time period 50 is about about 60% (.30 in top panel and .48 in bottom panel).

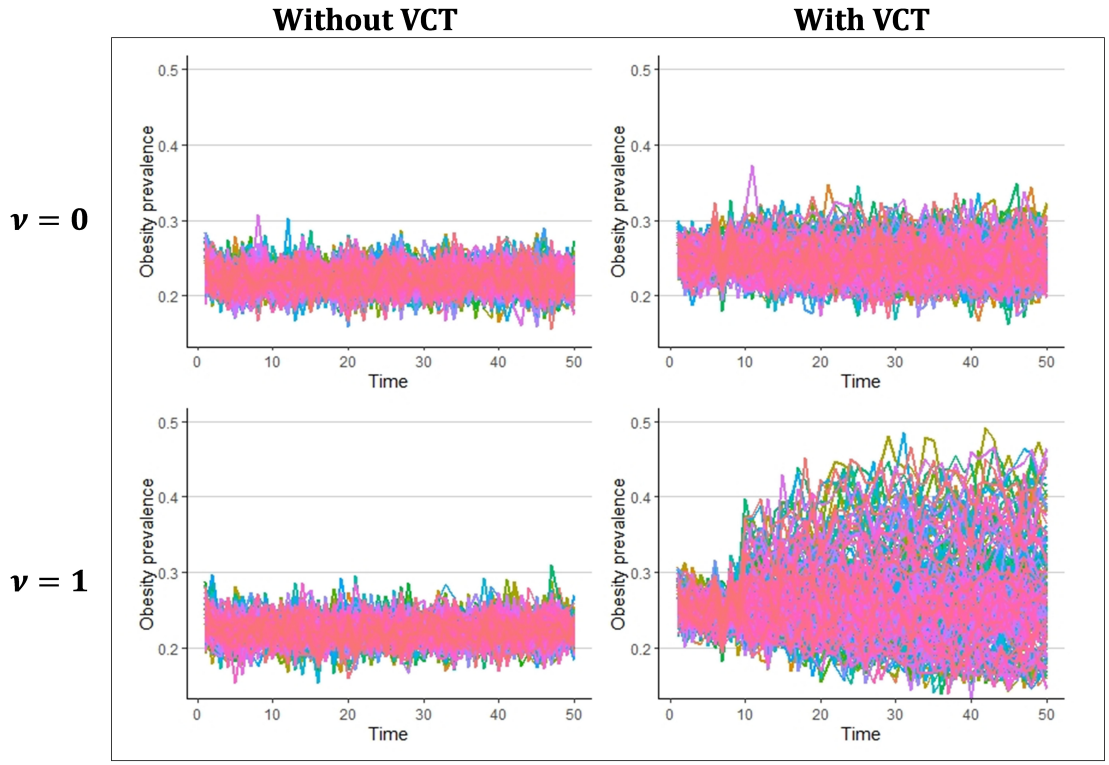


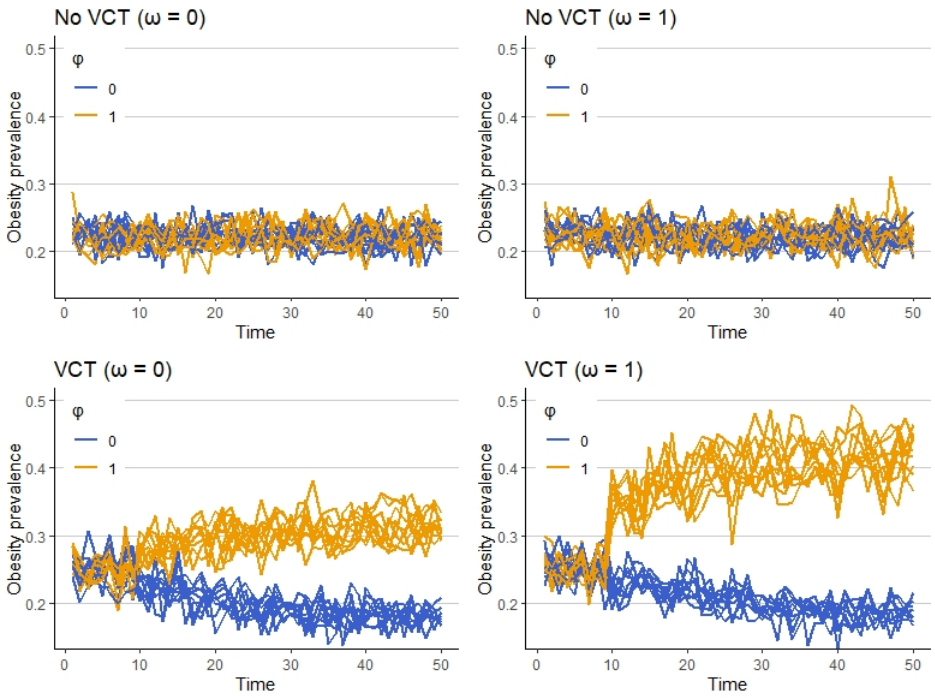
Figure 1: Time evolution of obesity prevalence at 40-44 years old in simulations without (left) and with (right) vertical cultural transmission, and with fidelity $\nu = 0$ (top) and $\nu = 1$ (bottom).

4.2 The impacts of DF and AM

How does the magnitude of the influence of AM compares to that of DF? Figure 2 suggests that when there is no VCT (top panel), contrasts are quite small. When VCT is fully included (bottom panel), important differences emerge. First, prevalence levels are always higher when DF is set to its maximum (yellow versus blue plots). Second, a shift from a fully random to a fully endogamous AM regime (left versus right plot) results in higher obesity prevalence regardless of DF. Third, the contrasts in prevalence between DF regimes are larger in a fully endogamous regime (yellow versus blue plots on right panel) than in a fully endogamous regime (yellow vs blue plot on left panel).

It follows that shifts in AM regimes can moderate to some extent the impact of DF on obesity prevalence but only when there is VCT (and high levels of fidelity)

To summarize: in the absence of VCT, both AM and DF play a minor role and differences in prevalence associated with minimum and maximum AM, on one hand, and minimum and maximum DF, on the other are small. Instead, when VCT is included larger contrasts emerge and the most important are associated with DF.



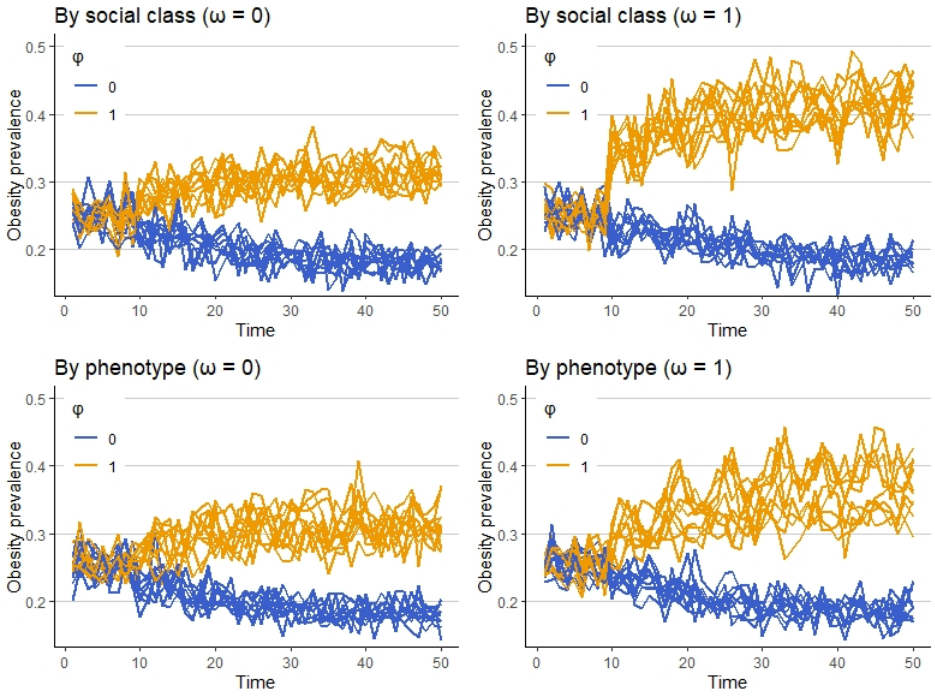
.

Figure 2: Time evolution of obesity prevalence at 40-44 years old in simulations with random mating (left) and fully endogamous mating by social class (right), when there is no differential fertility (blue) and when it is maximum (orange). The top panel shows simulations without VCT and the bottom panel with VCT. All curves shown have $\nu = 1$ and $a = 1$.

4.3 Does the nature of the AM regime matters?

Figure 3 displays phenotype trajectories under different DF regimes (blue vs yellow plots), by AM (left and right panels) and, finally, under different definitions of the AM regimes (top and bottom panels). Note that in both classes of AM, by social class and by phenotype, the fidelity of VCT is set to its maximum, e.g. $\nu = 1$, and so is $a = 1$. These plots show, first, that the AM regime does not

matter when mating is random (as it should, left panel). When mating is fully endogamous (right panel), social homogamy has a stronger effect on prevalence than phenotype endogamy. When AM is by social class, prevalence attains [.36 - .47] at $t = 50$ under differential fertility, while those are [.29 - .43] when AM is by phenotype (increases between 9 and 24 %).

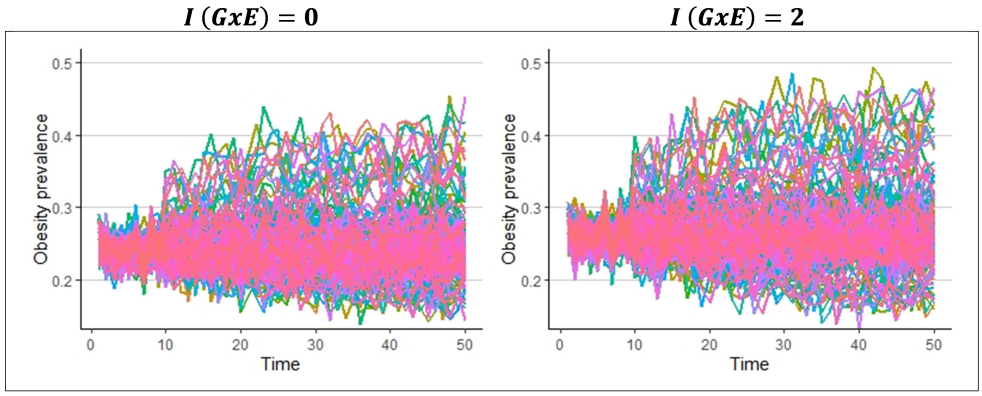


.

Figure 3: Time evolution of obesity prevalence at 40-44 years old in simulations with random mating (left) and fully endogamous mating (right), when there is no differential fertility (blue) and when it is maximum (orange). In the top panel, assortative mating is defined by social class. In the bottom panel, assortative mating is defined by phenotype (WHO-defined BMI class). All plots are from models where ν and a are set to 1.

4.4 Impact of GxE

Figure 4 displays obesity prevalence levels in two scenarios, one without GxE interaction (left panel) and the other with the effect of GxE set at the maximum blue we use in the simulation. As expected, GxE increases the prevalence levels in most cases. However, the contrasts are smaller that those associated with FD, AM and VCH.



.

Figure 4: Time evolution of obesity prevalence at 40-44 years old in simulations when the GxE term (see equation 2.2) is $I = 0$ (left) and $I = 2$ (right).

5 Discussion

Our results lead to a handful of potentially useful inferences. First, VCT plays a central role in the trajectory of obesity prevalence and its influence increases as fidelity of transmission increases. While the influence of GT is not trivial, it is overpowered by the impact of VCT.

Second, the role played by AM and DF is enhanced under conditions in which there is VCT. When fidelity is high, difference in prevalence by DF are larger than those associated with difference in AM. Furthermore, AM does exert a moderating, albeit modest, influence on the effect that DF has on the trajectories of obesity prevalence.

Third, when the class of AM is by social class rather than by phenotype, disparities in prevalence associated with DF and AM are increased. This result prompts two questions. One is whether the contrasts between classes of AM, by phenotype and by social class, are the same when ν is set at different levels lower than 1. The other question regards the pattern of results that is associated with a blended AM regime, e.g. by social class and phenotype simultaneously.

Fourth, even though we use moderate values of excess genetic penetrance (I) to reflect the impact of GxE interaction, its role is not trivial. In fact, it augments the impact of fertility differentials and through it, reinforces the power of potential interventions designed to eliminate or reduce exposure of some subgroups to obesogenic environments. In addition, its influence could be

felt in areas we did not fully explore. For example, a GxE interaction that emerges in generation G0 may have important effects on genotypic composition, couples' distribution by obesity category, and genetic heritability of the phenotype, that will be subsequently expressed in G1.

The paper has two shortcomings. First, although unlike other models, ours includes an empirically based representation of VCT and HCT, this is still highly stylized and does not capture the richness of their likely influence in real populations. Thus, although the CRS we employ is based on empirical findings regarding the influence of household domains on child body size, its estimation rests on two data sets from the US. These are unlikely to represent 'niche' conditions in populations at different stages of the obesity epidemic. By the same token, the influence of indirect genetic effects (IGE) is probably underestimated by the CRS as the genotypical information from parents and children is quite limited.

Second, the model only represents situations in which the relation between fertility and obesity is positive, e.g the observed pattern in populations that are in advanced stages of the obesity epidemic. A more realistic model should include both regimes simultaneously, one in which the initial stages are characterized by an inverse relation that is reversed once the population attains certain levels of obesity prevalence. This may turn out to be a powerful feedback mechanism that only an Agent-Based model can handle efficiently.

Despite these limitations, there is value in our contribution. Because we rely on empirically derived, not guessed, input parameters that represent well the relations between substrate processes included in the model, our assessment of their influence is empirically anchored and defensible.

References

[1] Fisher, Ronald A. The Genetical Theory of Natural Selection. New York, Dover, 2nd edition edition, 1958.

[2] Sewall Wright. Statistical genetics and evolution. Bulletin of the American Mathematical Society, 48(4):223–246, 1942.

[3] Crow, James F. and Kimura, Motoo. Introduction to Population Genetics. Frontiers in nutritional science. The Blackburn Press, Caldwell,NJ, 1970.

[4] Luigi L Cavalli-Sforza' and Marcus W Feldman. Cultural versus Biological Inheritance: Phenotypic Transmission from Parents to Children (A Theory of the Effect of Parental Phenotypes on Children's Phenotypes). page 20, 1973.

[5] Robert Boyd and Peter J. Richerson. Culture and the Evolutionary Process. University of Chicago Press, June 1988. Google-Books-ID: MBg4oBsCKU8C.

[6] Brian Charlesworth. Evolution in Age-Structured Populations (2nd Edition). Cambridge University Press, Cambridge, December 2009.

[7] Luigi Luca Cavalli-Sforza and Marcus W. Feldman. Cultural Transmission and Evolution. Princeton, N.J, 1981. 05311.

[8] Hao Shen and Marcus W. Feldman. Genetic nurturing, missing heritability, and causal analysis in genetic statistics. Proceedings of the National Academy of Sciences, 117(41):25646–25654, October 2020. Publisher: Proceedings of the National Academy of Sciences.

[9] Hao Shen and Marcus W. Feldman. Cultural versus biological inheritance: A retrospective view of Cavalli-Sforza and Feldman (1973). Human Population Genetics and Genomics, 1(1):0–0, October 2021. Publisher: Pivot Science Publication Corp.

[10] Hilla Behar and Marcus W. Feldman. Gene-culture coevolution under selection. Theoretical Population Biology, 121:33–44, May 2018.

[11] Nicole Creanza, Oren Kolodny, and Marcus W. Feldman. Cultural evolutionary theory: How culture evolves and why it matters. Proceedings of the National Academy of Sciences, 114(30):7782–7789, July 2017. 180 citations (Crossref) [2023-05-16].

[12] Marcus W. Feldman, Freddy B. Christiansen, and Sarah P. Otto. Gene-culture co-evolution: teaching, learning, and correlations between relatives. Israel Journal of Ecology and Evolution, 59(2):72–91, May 2013. 00005.

[13] Nestor Aldea-Ramos, Aitor Garcıa-Aguirre, Hiram Beltran-Sanchez, Sebastian Daza, and Alberto Palloni. Models of intergenerational transmission of body size.

[14] PD. Gluckman and PD. Hanson. The Develpmental Origins of Health and DIsease. Cambridge, 2006.

[15] Leslie O. Schulz and Lisa S. Chaudhari. High-Risk Populations: The Pimas of Arizona and Mexico. Current Obesity Reports, 4(1):92–98, March 2015.

[16] Stefan Walter, Iván Mejía-Guevara, Karol Estrada, Sze Y. Liu, and M. Maria Glymour. Association of a Genetic Risk Score With Body Mass Index Across Different Birth Cohorts. JAMA, 316(1):63, July 2016. tex.ids= walter2016b, walter2016c publisher: American Medical Association.

[17] Yiyue Huangfu, Alberto Palloni, Hiram Beltrán-Sánchez, and Mary C McEniry. Gene–environment interactions and the case of body mass index and obesity: How much do they matter? PNAS Nexus, 2(7):pgad213, July 2023. 0 citations (Crossref) [2023-08-31].

[18] Claude Bouchard. Gene–Environment Interactions in the Etiology of Obesity: Defining the Fundamentals. Obesity, 16(S3):S5–S10, 2008. 00000 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1038/oby.2008.528.

[19] Aitor Garcıa-Aguirre, Nestor Aldea, Mary McEniry, Yiyue Huangfu, Michael Lund, Thorkild Sørensen, and Alberto Palloni. Inherited Environments: Modeling the Cultural and Genetic Pathways of Childhood Obesity.

[20] Nicholas Christakis and James H Fowler. The spread of obesity in a large social network over 32 years. The New England journal of medicine, 357(4):370–9, July 2007. 2972 citations (Crossref) [2022-02-07] 00000.

[21] David A. Shoham, Ross Hammond, Hazhir Rahmandad, Youfa Wang, and Peter Hovmand. Modeling Social Norms and Social Influence in Obesity. Current Epidemiology Reports, 2(1):71–79, March 2015.

[22] Keisuke Ejima, Kazuyuki Aihara, and Hiroshi Nishiura. Modeling the obesity epidemic: social contagion and its implications for control. Theoretical Biology and Medical Modelling, 10(1):17, March 2013.

[23] John A. Dawson, Emily J. Dhurandhar, Ana I. Vazquez, Bo Peng, and David B. Allison. Propagation of Obesity Across Generations: The Roles of Differential Realized Fertility and Assortative Mating by Body Mass Index. Human heredity, 75(0):204–212, 2013. 00000.

[24] Benjamin W. Domingue, Jason Fletcher, Dalton Conley, and Jason D. Boardman. Genetic and educational assortative mating among US adults. Proceedings of the National Academy of Sciences, 111(22):7996–8000, June 2014. 130 citations (Crossref) [2023-05-16].

[25] Abdel Abdellaoui, Karin J. H. Verweij, and Brendan P. Zietsch. No evidence for genetic assortative mating beyond that due to population stratification. Proceedings of the National Academy of Sciences, 111(40), October 2014. 21 citations (Crossref) [2023-05-16].

[26] Dalton Conley, Thomas Laidley, Daniel W. Belsky, Jason M. Fletcher, Jason D. Boardman, and Benjamin W. Domingue. Assortative mating and differential fertility by phenotype and genotype across the 20th century. Proceedings of the National Academy of Sciences, 113(24):6647–6652, June 2016.

[27] Loic Yengo, Matthew R. Robinson, Matthew C. Keller, Kathryn E. Kemper, Yuanhao Yang, Maciej Trzaskowski, ..., and Peter M. Visscher. Imprint of assortative mating on the human genome. Nature Human Behaviour, 2(12):948–954, December 2018. 46 citations (Crossref) [2022-05-27] 00062.

[28] Philip W. Hedrick. Assortative Mating and Linkage Disequilibrium. G3: Genes|Genomes|Genetics, 7(1):55–62, October 2016. 6 citations (Crossref) [2022-02-07].

[29] Samuel H. Preston and Cameron Campbell. Differential Fertility and the Distribution of Traits: The Case of IQ. American Journal of Sociology, 98(5):997–1019, March 1993.

[30] Teresa A. Ajslev. Assortative marriages by body mass index have increased simultaneously with the obesity epidemic. Frontiers in Genetics, 3, 2012.

[31] Chika Vera Anekwe, Amber R. Jarrell, Matthew J. Townsend, Gabriela I. Gaudier, Julia M. Hiserodt, and Fatima Cody Stanford. Socioeconomics of Obesity. Current Obesity Reports, 9(3):272–279, September 2020.

[32] P. Jacobson, J. S. Torgerson, L. Sjostrom, and C. Bouchard. Spouse Resemblance in Body Mass Index: Effects on Adult Obesity Prevalence in the Offspring Generation. American Journal of Epidemiology, 165(1):101–108, October 2006. 41 citations (Crossref) [2022-05-27].

[33] Nathan Keyfitz. Introduction to the Mathematics of Population. Addison Wesley, 1968.

Supplemental Material

The code to reproduce the analyses performed in the paper is available in the repository `https://github.com/palloni/obesity-leslie-abm`.

1 Empirical justification of parameters

The value of input parameters used throughout are not educated guesses but reproduce closely what we found after an extensive search of empirical studies.

1.1 Assortative mating

The values of the input parameter ω partially reflects empirical values based on information from two sources. The first was the creation of artificial data set containing the distribution of BMI for 10,000 females. Their BMI was a normal variate with mean 27 and a standard deviation of 3. We then computed the BMI distribution of their 'partners' using predicted values from a regression with slopes, β, that varied between .10 to .70, normally distributed errors, ϵ with 0 and standard deviations ranging from 1 to 5. In all, we computed 35 different partners' BMI distributions. The R^2 values of the predictive regressions ranged from .004 to .81. In each case, we computed two additional quantities: (i) the χ^2 value of a cross-tabulation of females and their partners using the four obesity categories we have employed throughout and (b) the odds ratios, OR_{ij} of females in the i^{th} category of having partners in the j category.

With the artificial values of p_{ii} and π_i available from each of the 35 pairs of BMI distributions, we computed the value of ω using the expression

$$\omega = (p_{ii} - \pi_i)/(1 - \pi_i) \tag{S1}$$

which is simply derived from 2.3. We then estimate three regression models to sequentially express ω as a function of β, R^2 and χ^2.

The final step consists of searching the recent literature and identifying estimates of either *beta*, R^2, χ^2, or combinations thereof. Estimates of ω's consistent with a study's parameter(s) are

then calculated using one of the three regression models defined before.

Column 5 of Table S1 displays estimates of ω obtained from each of the studies identified in the first column using the parameter estimate in column 4. The range spanned by omega estimates is [.06-0.25] or the first 25 percent of the range of omega used in simulation. Thus, our inference regarding the *modest* direct role of AM shaping additive vertical genetic heredity (via increased LD or homozygosity) and, more generally, the global obesity epidemic, is probably an overstatement.

Table S1: Observed within-couple correlations of BMI/weight and omega

Study	**Population**	**Type of measure**	**Parameter estimate**	**Omega**
Ajslev et al. (2012)	37,792 pairs (Copenhangen)	BMI distributions	-	0 -.012
Allison et al. (1996)	Multiple studies	Couples' correlation weight	0.10-0.33	0.062-0.206
Hebebrand et al. (2000)	128-150 couples German National Nutritional Survey	BMI distributions	-	0.16-0.20
Katzmarzyck et al. (2002)	1341 parents Canadian Fitness Survey (1981)	BMI rank correlation	0.14	0.087
Sjaarda & Kutalik (2022)	51664 couples UK Biobank	Weight correlation	0.25	0.155
Speakman et al. (2007)	42 couples Northeast Scotland	BMI	0.33	0.206
Authors' estimate from HRS	1,568 couples in 2006 wave	BMI distributions	-	0.06
Authors' estimate from HRS	All couples - all waves (up to 2020)	BMI distributions	-	0.11
Authors' estimate from DHS India 2019-20	38,857 couples	BMI distributions	-	0.12

Estimates from Allison et al (1996) correspond to correlation of partners' weight (not BMI). The range of values (0.10-0.33) includes Allison's et al. own and those and those from 29 different studies in sub-populations from USA, UK, Italy, Sweden, Norway, Denmark, Brazil, Peru, Israel (see Table 1 in Allison et al. (1996)).

1.2 Differential fertility

To establish an empirically plausible range for the differential fertility parameter ω we relied on the observed relation between maternal obesity and number of children ever born (controlling for maternal age). We used information for multiple sources: 33 African Demographic Health Surveys, the US National Longitudinal Survey of Youth (NLSY) as well as the National Health and Nutrition Examination Survey (NHANES). The populations included in our surveys represent a very broad range of countries at different stages of the obesity epidemic, from those in which it is not yet discernible to those that have attained relatively high values (though not the highest). Admittedly, this is not an ideal data set because in all cases it lacks information on the father or partner. But, alas, there are no national data that include anthropometry of both members of a couple couple. To approximate the values of the parameter φ we first estimated regressions of the number of children ever born using a dummy variable for obesity and controls (age and education levels).

The estimated effect of the dummy variable, β and φ are related by the following expressions:

$$\beta_{max} = \varphi \times 3 \tag{S2}$$

$$\beta_{min} = \varphi \times .5 \tag{S3}$$

where β_{min} and β_{max} are the minimum and maximum values consistent with a value of φ.

Table S2 displays values of β from sample surveys of some populations and subpopulations. The figures in this table confirm that the value of φ in the middle of the range we are using is consistent with minimum and maximum β values of .25 and 1.5 respectively. Because the range of β values is approximately [.20,.50], they are consistent with φ in the range [.067, 1].[12].

An important feature of the table is the association between the magnitude of β's and the population prevalence of obesity (last column). In particular, there is a strong positive relation (close to that observed in the US) among females with highest education in countries with the lowest prevalence of obesity (Africa, India). This is consistent with the idea that as the obesity

[12]The maximum and minimum values were computed using Table **??**

Table S2: Estimates of Relations between Children Ever Born and Maternal Obesity

Country	Source	Population	β estimate	Adult Female Prevalence
India	DHS	All	-0.381	0.07
		Low Ed	-0.181	
		High Ed	0.05	
Turkey	DHS	All	0.604	0.41
		Low Ed	0.054	
		High Ed	0.496	
Africa[1]	DHS	All	-0.792	0.017-0.150
		Low Ed	-0.371	
		High Ed	0.206	
Asia[1]	DHS	All	-0.157	0.010-0.044
		Low Ed	-0.237	
		High Ed	0.168	
USA	NLSY 1997	All	0.246	0.37
	NHANES	All	0.323	0.37

[1.] Africa includes 33 DHS samples and Asia 8 DHS samples.
[2.] All regression coefficients are significant at $p < 0.001$.
[3.] Source of obesity prevalence estimates: `https://ncdrisc.org/obesity-prevalence-ranking.html`

epidemic advances, there is a transition from a negative relation between obesity and fertility to a positive one.

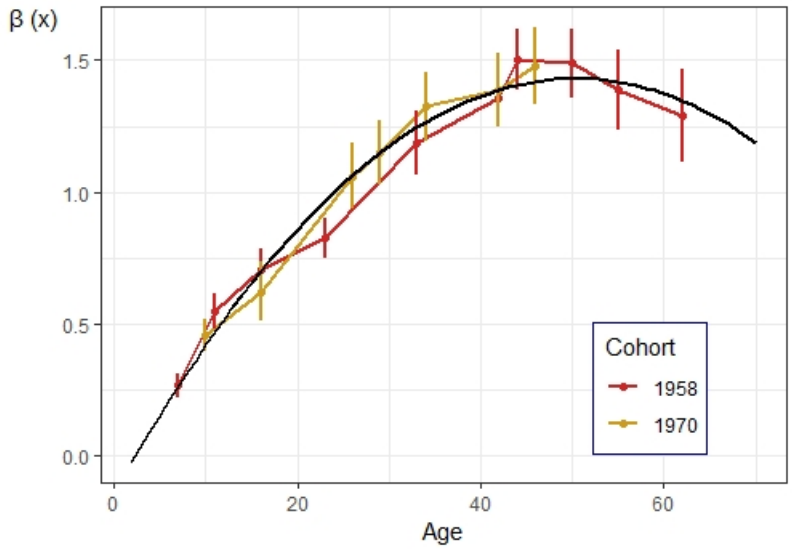
1.3 Genetic penetrance, baseline BMI

The genetic penetrance in equation (3.1) of the main text, $\beta(x)$, as well as the baseline BMI, $\alpha(x)$, are derived from the 1958 National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70) in the United Kingdom (England, Scotland and Wales)[13]. The follow-up of this cohort includes BMI measures or reportings at various ages, from birth to 62 years old in the NCDS and from birth to 46 years old in the BCS70. Polygenic Risk Scores for BMI are computed for a large part of the samples.

Figure S1 displays the effect of the PRS on BMI for both cohort at succesive ages, drawn from linear regressions. As, for the simulation, $\beta(x)$ values are needed at all age classes (not only those present in the cohorts), we fitted the obtained results to a quadratic curve in order to obtain

[13]We accessed the restricted NCDS and BCS70 data after a request procedure to the Center for Longitudinal Studies, University College London.

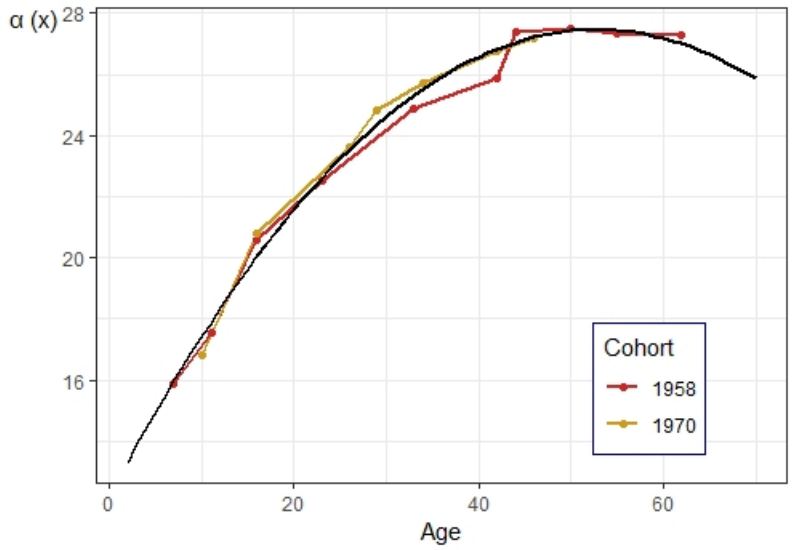
estimates for all ages. For the simulation, the value at the middle of the age class is considered (e.g. $\beta(22.5)$ for age class 20-24). Note than the maximum of the fitted curve for $\beta(x)$ occurs at age 50.



.

Figure S1: Effect of PRS on BMI by age in the 1958 NCDS and the 1970 BCS, with 95 % confidence intervals, and quadratic regression line.

The same procedure is used to retrieve $\alpha(x)$ values. Figure S2 shows the values for the intercept of the regression of BMI on PRS for both cohorts, and the fitted quadratic regression line.Here the maximum occurs at 53 years old.



.

Figure S2: Intercept for the regression of BMI on PRS by age in the 1958 NCDS and the 1970 BCS, and quadratic regression line.

2 Supplementary results

Table S3: Coefficients of regression of $\log(BMI)$ for children and adults, in simulations with and without vertical cultural transmission, at time $t = 20$

	10-14 years old		**40-44 years old**	
	VCT	**No VCT**	**VCT**	**No VCT**
Const.	2.895***	2.920***	3.258***	3.290***
$\beta(\omega)$	0.003	-0.001	**0.009***	0.003
$\beta(\varphi)$	**0.019*****	0.000	**0.027*****	0.002
$\beta(I(GxE))$	**0.006*****	0.001	**0.005****	0.001
$\beta(\nu)$	**0.015*****	0.000	0.003	0.000
$\beta(a)$	0.000	0.001	**-0.010*****	0.001
$\beta(\omega \times \varphi)$	0.010	0.003	0.002	-0.006
$\beta(\omega \times I(GxE))$	-0.001	0.000	-0.003	-0.001
$\beta(\varphi \times I(GxE))$	0.004	-0.001	-0.003	-0.001
$\beta(\nu \times a)$	**0.011*****	0.001	**0.035*****	0.001
$\beta(\omega \times \varphi \times I(GxE))$	0.001	0.000	0.006	0.002
N	360	360	360	360
R^2	0.682	0.041	0.669	0.016
Adj. R^2	0.673	0.013	0.660	-0.012
RSE	0.012	0.006	0.014	0.007
df	349	349	349	349

The N for corresponds to 72 combinations of parameter values and 5 relicas for each of them.