Understanding variation in health-related behaviors: Evidence from American twins

Nikkil Sudharsanan, Hans-Peter Kohler, Jere R. Behrman

University of Pennsylvania

Abstract

Health-related behaviors contribute significantly to U.S. morbidity and mortality, yet empirical evidence on causes of within-population variation in health-related behaviors is mixed. For example, schooling is frequently assumed to affect different health-related behaviors, but several studies find that schooling does not explain much variation in health behaviors. This paper advances knowledge on the relative importance of schooling, genetic endowments, and environments using a novel ACE-β twin model that integrates behavioral-genetic approaches focused on variance decomposition and economics within-MZ twin pair design focused on causal inferences. We find that schooling does not have significant effects on health-related behaviors. A large proportion of variation in BMI and health seeking behavior is due to genetic endowments with a negligible contribution from common environments. This pattern is reversed for cigarette smoking with common environments contributing a large portion to the overall variation. We conclude by discussing the policy implications of our findings.

Extended Abstract

Introduction

Unhealthy behaviors, such as smoking and heavy drinking, are responsible for a large portion of disease and deaths in the United States. For example, McGinnis and Foege estimated that smoking, heavy drinking, and obesity were responsible for 38% of United States mortality in 1993; for the year 2000, Mokdad and coauthors attribute almost 50% of mortality in the United States to behaviors, with smoking, weight, and alcohol consumption as the primary contributors.[1,2] Unhealthy behaviors extend beyond smoking, drinking, and physical inactivity, and can be defined as "any action, or deliberate inaction, by an individual that affects [their] own health".[3]

Given the apparent substantial contributions of behaviors to health and mortality, a large literature has focused on why people engage in behaviors that are widely known to negatively affect health. It is well-known that the extent to which individuals engage in unhealthy behaviors varies widely within societies and institutional contexts.[4–6] For example, in many studies, socioeconomic status, usually measured as either schooling or household income, is a strong predictor of health-related behaviors. As Cawley and Ruhm show, "compared to high school dropouts, college graduates [in the U.S.] were 13.9 percentage points less likely to smoke, 8.7 percentage points less likely to be obese, 0.9 percentage points less likely to drink heavily, and 22.3 percentage points less likely to be physically inactive. In addition, they are 12.6 percentage points more likely to receive mammograms, 15.3 percentage points more likely to receive colorectal screening and 16.0 percentage points more likely to use sunscreen when outside on warm sunny days".[3] Higher levels of schooling are overwhelmingly associated with healthier behaviors across many domains. The positive associations between schooling and health-related behaviors can thus potentially explain why more-educated people tend to be in better health. Indeed, in the United States, there are large educational differences in smoking, obesity, and heavy drinking.[4]

Economic studies of the underlying behavioral causes of health outcomes often fail to explain a large portion of the variation in individuals' health-related behaviors. Most economic studies of health behaviors are influenced by Grossman's model of health capital, in which education affects health because more-educated people are more likely to make better choices regarding health inputs, including health-related behaviors, given available resources including time (allocative efficiency), or are better at producing health from a given set of inputs (productive efficiency).[7] Cutler and Glaeser try to empirically confirm Grossman's model by arguing that if health-related behaviors are determined by individual investments in future health, different health-related behaviors should be correlated within individuals. However, using data from the Behavioral Risk Factor Surveillance System, they find weak correlations among the health behaviors of individuals—such as obesity and smoking, and smoking and receiving mammograms for women—implying that the factors that determine health-related behaviors vary across behavioral domains (e.g. the factors that lead individuals to smoke do not necessarily lead individuals to be physically inactive).[8] Moreover, despite significant correlations between schooling and behaviors, schooling explains very little of the total variation in health-related behaviors.[4]

Variation in health-related behaviors has also been examined from a behavioral genetic perspective. For example, one study in the United States found that BMI has a heritability of 72%, implying – subject to the assumptions of the behavioral genetic model – that 72% of the within-population variation in BMI can be attributed to variation in genetic factors. Exercise has an estimated heritability of 26%, smoking 30%, and heavy drinking 38%.[8] In all cases, therefore, genetic heterogeneity seems to be an important factor contributing to variation in health-related behaviors. Similarly, other studies have argued that childhood obesity is importantly related to unobserved genetic endowments, the family environment, and the interaction of genes and family environmental factors "working in concert".[9] A limitation of these behavioral genetic studies is that they estimate the heritability of health behaviors, but do not consider potential pathways between schooling and health that are a focus of economic models.

Twin studies provide a way to reconcile both the economic and behavioral genetic models of health-related behaviors. In economic approaches to twin studies, schooling and health are assumed to be jointly determined by unobserved genetic and environmental endowments; therefore, comparing differences between twins can shed light on the causes of health. Fixed-effects twin studies estimate the causal effect of schooling on health by looking at within-twin differences in schooling and health behavior. Identical twins are assumed to share identical genetic endowments at conception and substantial childhood environments, so comparing differences between two identical twins purges the schooling effect of bias due to important unobserved characteristics related to both schooling and health. Within-MZ twins studies, for example, have recently shown that the causal contributions of schooling to health are substantially smaller than the cross-sectional correlations between schooling and health. [10–12]

Standard behavioral genetic "ACE" models investigate the health-schooling relationship by assuming that schooling and health are both the result of genetic, common environmental ("shared environments" between twins), and individual-specific factors. These studies then estimate the fraction of variance in schooling and health behavior explained by genetics, common environments, and individual environments, and they identify the extent to which these factors jointly affect both schooling and health. Importantly, however, these models do not estimate a direct causal effect of schooling on health behavior, i.e., the effect that is central to the economic twins-approach to the study of the health-schooling relationship. Instead, standard ACE models assume that the relationship between schooling and health behavior is only due to their joint correlation with genetics, common environments, and individual endowments. Examples of these studies are [13,14].

The goal of this paper is to bridge the economic and behavioral genetic perspectives to better understand the underlying sources of variation in health behaviors. We accomplish this by using a novel ACE-β twin model that both identifies the causal effect of schooling and estimates the contribution of genetics, common environments, and individual environments to the total variation in health behaviors.[15]

Data

Our analyses for the final paper will use four well-known datasets of American twins. Our current preliminary analyses use the National Survey of Midlife Development in the United States II survey and data from the Socioeconomic Survey of Twins of the Minnesota Twin

Registry; however, our final analyses will also incorporate the National Longitudinal Study of Adolescent to Adult Health and the Mid-Atlantic Twin Registry data sets.

National Survey of Midlife Development in the United States (MIDUS II)

MIDUS II is a sample of 1,484 twins in the United States aged 35-85 collected between 2005 and 2006. The MIDUS II includes a rich set of health data on twins; however, all results are self-reported and may be subject to reporting bias.

Based on a separate zygocity survey given to MIDUS respondents identified as twins, twins were classified as monozygotic (MZ), dizygotic (DZ) of the same sex, or dizygotic with different sexes. We use a continuous measure of age, calculated by the MIDUS II researchers using self-reported date of birth and date of survey. We do not consider the race of individuals in our analyses since the sample was mostly white (1,331 of the 1,484 individuals).

Schooling was categorically reported; we assigned completed grades of schooling to individuals as follows: No school/some grade school (3 grades), eighth grade/junior high school (7 grades), some high school (10 grades), GED (10 grades), graduated from high school (12 grades), 1 to 2 years of college (13 grades), graduated from a 2-year college (14 grades), 3 or more years of college (15 grades), graduated from a 4- or 5-year college (16 grades), some graduate school (17 years), master's degree (18 grades), doctoral degree (21 grades). We use three measures of health behaviors: Body mass index (BMI) as a measure of diet and exercise behavior, cigarettes smoked per day, and number of routine health visits per year as a measure of health-seeking behavior. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared. Individuals were asked how many cigarettes a day they smoked during their heaviest year of smoking; based on this question we created a continuous variable of the number of cigarettes smoked per day (never-smokers were coded as zero). We created a continuous variable measuring the number of times on average an individual saw a doctor per year for routine physical exams.

Of the 1,484 individuals in the sample, 238 individuals were dropped because their co-twin was not in the sample. 78 individuals were dropped because they were triplets or greater. 476 individuals were dropped due to missing information. Because analysis of mixed-sex pairs confounds biological and behavioral factors, 150 opposite sex DZ twins were dropped. 12 individuals were dropped due to missing zygocity information. Finally 1 twin pair (2 individuals) was dropped because one of the twins was pregnant. The final sample consisted of 528 individuals (148 MZ twin pairs and 116 DZ twin pairs).

The Minnesota Twin Registry (MTR)

The MTR is a registry of all twins born between 1936 and 1955 in Minnesota. Our data are from the Socioeconomic Survey of Twins, a survey of 3,631 same-sex MZ and DZ twins conducted in 1994. We approximated the age of individuals by subtracting the individual's birth year from 1994. Individuals were asked to report the highest grade of 1^{st}-12^{th} grade school they had completed and if they had completed a vocational, associate, bachelor, masters, or doctoral degree. Based on these self-reports we created a continuous measure of completed grades of schooling. We calculated BMI from self-reported height and weight. Unfortunately, data on cigarette smoking or health-seeking behaviors were not available.

Of the 3,631 individuals, 833 individuals were dropped because their co-twin was not in the sample. Additionally 224 individuals were dropped due to missing schooling or BMI data. Our final sample contained 2,350 individuals (659 MZ twin pairs and 516 DZ twin pairs).

Methods

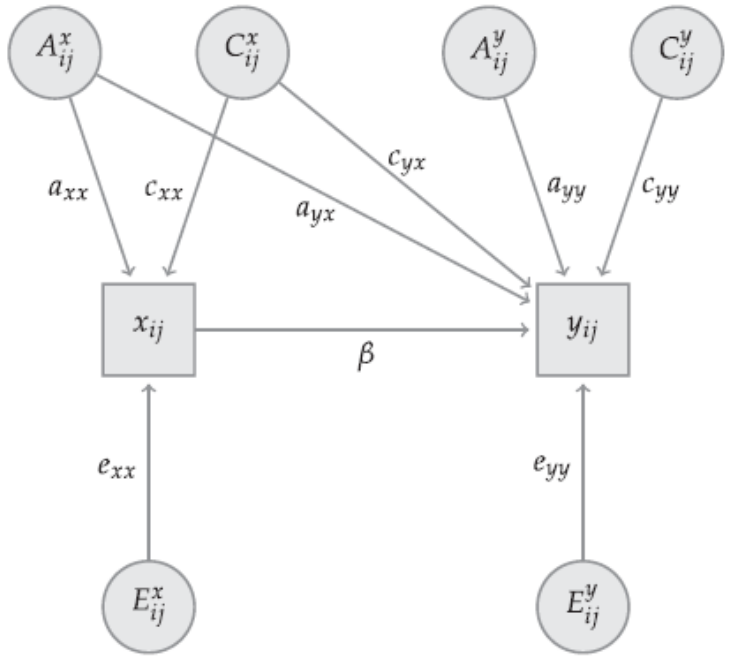
We assume that schooling is produced from the additive effects of genetic endowments, the common (mostly childhood) environment shared by members of a twinship, and individual idiosyncratic environments. Therefore, for twin *i* in pair *j*, schooling can be represented as:

$$x_{ij} = a_{xx}A_{ij}^{x} + c_{xx}C_{j}^{x} + e_{xx}E_{ij}^{x} .$$

Here x_{ij} is schooling attainment, A_{ij}^{x}represents genetic endowments, C_{j}^{x} is the common environment, and E_{ij}^{x} are individual environments. We then assume that schooling has a direct additive effect on health-related behavior y_{ij}, which is itself influenced by individual idiosyncratic environments E_{ij}^{y}, common environmental factors C_{j}^{y}, and genetic endowments specific to the health-related behaviors A_{ij}^{y}. We also allow health-related behaviors to be affected by the environmental and genetic endowments that affect schooling, C_{j}^{x} and A_{ij}^{x} respectively. This results in the following specification:

$$y_{ij} = \beta(x_{ij}) + a_{yx}A_{ij}^{x} + c_{yx}C_{j}^{x} + a_{yy}A_{ij}^{y} + c_{yy}C_{j}^{y} + e_{yy}E_{ij}^{y}.$$

Visually, this specification can be represented by the following path diagram:



To estimate all nine parameters of the model, we stack observed phenotypes *xij* and *yij* for twin 1 and twin 2 in twin pair j in vectors:

$P_{mz} = (x_{1j}, y_{1j}, x_{2j}, y_{2j})'$ for MZ twins,

$P_{dz} = (x_{1j}, y_{1j}, x_{2j}, y_{2j})'$ for DZ twins.

The expected variance-covariance matrices for both MZ and DZ twins can be expressed as the expectation of PP'. We impose the additional constraints that MZ twins share identical genetic endowments and common environments, DZ twins share common environments and 50% of their genetic endowments (i.e., no assortative mating for the twins' parents), and individual environments related to schooling do not affect health behavior. Formally:

$cov(A_{1j}^{MZ}, A_{2j}^{MZ}) = 1,$

$cov(C_{1j}^{MZ}, C_{2j}^{MZ}) = 1,$

$cov(A_{1j}^{DZ}, A_{2j}^{DZ}) = 0.5,$

$cov(C_{1j}^{DZ}, C_{2j}^{DZ}) = 1,$

$e_{xy} = 0.$

Finally, we standardize A, C, E to have means of zero and variances of one. This set-up corresponds to nine free parameters and nine unique moment conditions, exactly identifying each parameter. We used a maximum likelihood algorithm to estimate the nine parameters by minimizing the difference between the observed and expected variance-covariance matrices.

Using the parameters, the total variation in schooling and health behavior can be expressed as:

$\sigma_x^2 = (a_{xx}^2 + c_{xx}^2 + e_{xx}^2),$

$\sigma_y^2 = \beta^2\sigma_x^2 + 2\beta(a_{xx}a_{yx} + c_{xx}c_{yx}) + (a_{yy}^2 + a_{yx}^2 + c_{yy}^2 + c_{yx}^2 + e_{yy}^2).$

The fractions of the variance explained by genetics, common environments, and individual environments are then calculated as

$A_{std} = \frac{\beta^2 a_{xx}^2 + 2\beta(a_{xx}a_{yx}) + (a_{yy}^2 + a_{yx}^2)}{\sigma_y^2},$

$C_{std} = \frac{\beta^2 c_{xx}^2 + 2\beta(c_{xx}c_{yx}) + (c_{yy}^2 + c_{yx}^2)}{\sigma_y^2},$

$E_{std} = \frac{\beta^2 e_{xx}^2 + e_{yy}^2}{\sigma_y^2}.$

For this analysis, we first show descriptive statistics for the two twin samples, including the within-twin differences for each variable. Next, we present the parameter estimates from this twin model and the decomposition of the variance in health behaviors. Analyses were conducted in Stata 12 and R 3.0.3 using the OpenMx package.

Twins studies have been criticized on methodological grounds, but most of these criticisms are more applicable to heritability coefficients than to using twins as controls to attempt to ascertain causal effects, e.g. of neighborhoods on mental health. It has been noted, for example, that MZs

are not perfectly identical genetically, especially when considering epigenetic processes.[16] Although such considerations mean that the control for unobservable factors afforded by MZs is less than it would be if they also controlled for epigenetic processes, they do not negate the substantial advantages of twins controls over uncontrolled population-based studies that simply ignore genetic processes and unobserved childhood family background characteristics in exploring associations between risks and outcomes. Similarly, the validity of the so-called equal environment assumption, which holds that MZs share no more common environmental experiences than DZs, has been questioned.[17] Nevertheless, this hypothesis is testable and has generally been supported in the literature;[18] once again, such criticism is more relevant to computation of heritability coefficients than to the analyses proposed here of estimating the social determinants of health behaviors (although, since our ACE-β model is built on the behavioral genetic model, some of these criticism apply, and we will in our final version also estimate within-MZ analyses that are more robust with respect to this assumption). Yet another criticism holds that modern genomic methods and detailed biological understanding of genomics have caused twins-based methods to become antiquated.[19] We think, to the contrary, as Genome Wide Association Studies (GWAS) the often identify only very small single-gene effects on health and behaviors identified, and twin and related study designs continue to be relevant to obtain a comprehensive assessment of the genetic and social determinants of health and health-related behaviors. Finally, it has been questioned whether twins samples are representative of the populations from which they were drawn. Once again, this hypothesis is testable, and studies have generally reported little or no differences between twins and singleton populations. For example, a recent study that performed MRI brain scans[20] found no significant differences between twins and unrelated, age- and sex-matched singletons in several brain structures.

Results

The first column of Table 1 presents the means and standard deviations of age, sex (percent of the sample that is male), and the primary health-related behaviors variables. The Minnesota twins are younger on average (47 vs. 53 years), with a smaller percentage of men (35% vs. 39% male). Both samples have an average schooling attainment of about 14 years. The body mass index is also comparable between the Minnesota and MIDUSII twins (25.84 vs. 27.38) with the average of both groups being in the "overweight" category. For the MIDUSII twins, the average number of cigarettes smoked a day during the heaviest year of smoking is 5.05, but there is a large variation around this number (SD=11.07). Finally individuals in the MIDUSII sample reported on average 1.7 routine health visits per year.

Because our model estimates parameters for both MZ and DZ twins, our results might be biased if the MZ and DZ twin samples were drastically different on observed characteristics. Based on columns two and three of Table 1, we find that MZ and DZ twins are very similar across all the observed variables.

Figure 1 shows the distribution of within-twin differences for each of the variables. Because our parameters are identified using within twin-variation, our estimates do not represent a true average treatment effect (ATE) but rather a local average treatment effect (LATE) over the range of observed within-twin differences. For both samples, the modal twin difference in schooling is zero; however, some twins differ by as much as six years of schooling. We also find a fairly wide

distribution of within-twin differences in BMI in both samples. For the MIDUSII twins, we also observe wide ranges for the within-twin differences in cigarettes per day and health visits per year. Although our effects are not true ATEs, they are identified over a fairly wide range of within-twin differences; this is an improvement over studies in the literature that use minimum schooling-leaving age or "natural policy experiments" to estimate LATEs, because such estimates are only generalizable to the small window of the population on the margin of the policy change.[21–23]

Table 2 presents the effects of schooling, genetics, shared environments, and individual environments on health-related behaviors for the two twin samples. Each coefficient has been standardized to a z score so that the magnitudes can be directly compared. Table 3 then decomposes the overall variance in health-related behaviors into the fraction explained by genetics, common environments, and individual-specific environments. The coefficients a_{xx}, c_{xx}, and e_{xx} measure the influence of these latent variables on schooling, which in turn might affect health through the direct effect of schooling on health (β). The first important conclusion is that across both twin samples for every outcome, we do not find causal effects of schooling on health-related behaviors, and for none of the outcomes and datasets in Table 2 is the coefficient β significantly different from zero. This finding agrees with other studies that have found no or only very small causal effects of schooling on health after controlling for endowments. [11,12,15] The magnitudes of the estimated schooling effects are also very small compared to some of the other parameters.

We find that genetic endowments provide the largest contribution to variation in BMI in both datasets. Genetics also explains a large fraction of the variance health visits per year. Interestingly, genetic endowments do not contribute in a statistically significant way to variation in cigarettes per day; primarily individual-specific and, secondly, common environments explain the greatest change in cigarettes per day. Based on the results in Table 3, genetic endowments explain the largest portion of the variance in BMI (60.8% for the MIDUSII Twins and 67.9% for the Minnesota Twins) with almost no contribution from common environments. For cigarettes smoked per day, individual environments are the single largest contributor to overall variance (53.5%) with common environments as the second largest component (43.6%), and a very small contribution from genetics (2.9%). Most of the variance in health visits per year is due to individual environments (73.4%) but genetic variation does contribute a quarter of the overall variation.

Across all models, we find that unique individual variation has large statistically significant effects on health-related behaviors and explains large portions of the overall variation in such behaviors. This implies that individual characteristics uncorrelated between twins and unrelated to genetics and the common environment are responsible for large changes in health-related behaviors.

Discussion and conclusion

Health-related behaviors are significant contributors to morbidity and mortality in the United States, yet the empirical evidence on the underlying causes of the vast within-population variation in health-related behaviors is mixed. Schooling frequently is assumed to be a primary cause for different health-related behaviors, although some empirical work has found that schooling does not explain much of the overall variation in health behaviors. Other work has

focused on the contribution of genetic endowments, the childhood environment, and gene-environment interactions to variation in health behavior. This paper seeks to advance our knowledge on the relative importance of schooling, genetic endowments, and the childhood environment using an ACE- β twin model.

Our first primary conclusion is that across both datasets examined to date, schooling does not appear to have substantial or significant effects on health-related behaviors. This is consistent with some prior twin-studies that use the economics fixed-effects approach, including [11,12,15]. The estimates from this paper differ from economic studies of the effect of schooling that use natural experiments and instrumental variables.[21–23] Although most of these studies find that schooling has a plausibly causal effect on health, these results are only identified for very specific margins of the population, and thus are usually not generalizable to larger populations. If we assume that twins do not differ significantly from non-twins with respect to schooling or BMI, these results are identified for a larger subset of the population and come closer to estimating an average treatment effect.

Beyond schooling, we find that a very large proportion of variation in BMI is due to genetics with a negligible contribution from common environments. This pattern is reversed for cigarette smoking with common environments contributing a large portion to the overall variation. This conclusion is consistent with research that finds early exposure to cigarettes to be a strong predictor of smoking in adulthood.[24,25] Therefore, modifying the common shared childhood environment may be a plausible policy option to reduce smoking in adulthood. Although genetics cannot be altered as a policy solution and common environments contribute negligibly to BMI and health-seeking behavior, large fractions of the variance of both these behaviors are due to individual environments. By identifying additional sources of variation for these behaviors, future research may be able to generate feasible policy solutions.

Limitations

There are a few identification assumptions needed for the ACE- β schooling estimates to be causal: (1) Individual environments are assumed to be uncorrelated with the outcome; that is, there cannot be some unobserved factor aside from genetic endowments and common environments that affects both schooling attainment and the health-related behavioral outcome, though the violation of this condition produces predictable bounds on the causal estimates (see: [15]). (2) The outcome variable for one twin cannot depend on the outcome variable for another twin beyond their joint dependence on genetic endowments and childhood environments, although the violation of this condition produces predictable biases (see: [15]). For our estimates of the variance attributable to common environments, we also assume that the common environments of MZ twins are the same as the common environment of DZ twins. After controlling for any unobserved difference between twins through the within-twin estimates, we assume that the population of twins is representative of the larger American population and that the underlying causes of schooling and BMI are the same for twins as for the American population. The descriptive results suggest that the twin populations are reasonably representative of the *white* American population, although the twin populations tended to have slightly more schooling. The sample is overwhelming white, and the results estimated might not be generalizable to the unique childhood contexts experienced by other race/ethnic groups in the United States or in other societies.

Next Steps

Our planned next steps are:

1. Estimate the model for two additional twin datasets.
2. Determine if we can improve the model fit by modifying the specification or relaxing some of the assumptions.
3. Explore the use of this model for dichotomously coded health-related behaviors such as smoker/non-smoker.

References

1. McGinnis J, Foege W. Acutal Causes of Death in the United States. *JAMA J Am Med Assoc*. 1993;270:2207–2212.

2. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA*. 2004;291:1238–1245. doi:10.1001/jama.293.3.293.

3. Cawley J, Ruhm CJ. The Economics of Risky Health Behaviors. *Handb Heal Econ*. 2011;2:95–199.

4. Cutler D, Lleras-Muney A. Understanding Differences in Health Behaviors by Education. *J Health Econ*. 2010;29(1):1–28. doi:10.1016/j.jhealeco.2009.10.003.Understanding.

5. Case A, Menendez A. Sex differences in obesity rates in poor countries: evidence from South Africa. *Econ Hum Biol*. 2009;7(3):271–82. doi:10.1016/j.ehb.2009.07.002.

6. Banks J, Marmot M, Oldfield Z, Smith JP. Disease and disadvantage in the United States and in England. *JAMA*. 2006;295(17):2037–45. doi:10.1001/jama.295.17.2037.

7. Grossman M. On the concept of Health Capital and the Demand for Health. *J Polit Econ*. 1972. Available at: http://www.ppge.ufrgs.br/GIACOMO/arquivos/eco02072/grossman-1972.pdf. Accessed April 23, 2014.

8. Cutler D, Glaeser E. What Explains Differences in Smoking, Drinking, and Other Health-Related Behaviors? *Am Econ Rev*. 2005:238–242. Available at: http://www.nber.org/papers/w11100. Accessed April 7, 2014.

9. Birch L, Davison K. Family environmental factors influencing the developing behavioral controls of food intake and childhood overweight. *Pediatr Clin North Am*. 2001:1–11. doi:10.1016/S0031-3955.

10. Amin V, Behrman JR, Kohler HP. Schooling has smaller or insignificant effects on adult health in the US than suggested by cross-sectional associations: New estimates using relatively large samples of identical twins. *Social Science and Medicine*. 2014.

11. Behrman JR, Kohler H-P, Jensen VM, et al. Does more schooling reduce hospitalization and delay mortality? New evidence based on Danish twins. *Demography*. 2011;48(4):1347–75. doi:10.1007/s13524-011-0052-1.

12. Fujiwara T, Kawachi I. Is education causally related to better health? A twin fixed-effect study in the USA. *Int J Epidemiol*. 2009;38(5):1310–22. doi:10.1093/ije/dyp226.

13. McCaffery JM, Papandonatos GD, Lyons MJ, Koenen KC, Tsuang MT, Niaura R. Educational attainment, smoking initiation and lifetime nicotine dependence among male Vietnam-era twins. *Psychol Med*. 2008;38:1287–1297. doi:10.1017/S0033291707001882.

14. Tholin S, Rasmussen F, Tynelius P, Karlsson J. Genetic and environmental influences on eating behavior: The Swedish Young Male Twins Study. *Am J Clin Nutr*. 2005;81:564–569. doi:81/3/564 [pii].

15. Kohler H-P, Behrman J, Schnittker J. Social Science Methods for Twins Data: Integrating Causality, Endowments, and Heritability. *Biodemography Soc Biol*. 2011;57(1):88–141. doi:10.1080/19485565.2011.580619.

16. Petronis A. Epigenetics and twins: three variations on the theme. *Trends Genet*. 2006;22:347–350. doi:10.1016/j.tig.2006.04.010.

17. Brown BB, Yamada I, Smith KR, Zick CD, Kowaleski-Jones L, Fan JX. Mixed land use and walkability: Variations in land use measures and relationships with BMI, overweight, and obesity. *Heal Place*. 2009;15:1130–1141. doi:10.1016/j.healthplace.2009.06.008.

18. Joseph J. The equal environment assumption of the classical twin method: A critical analysis. *J Mind Behav*. 1998;19:325–358.

19. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. A test of the equal-environment assumption in twin studies of psychiatric illness. *Behav Genet*. 1993;23:21–27. doi:10.1007/BF01067551.

20. Ordaz SJ, Lenroot RK, Wallace GL, et al. Are there differences in brain morphometry between twins and unrelated singletons? A pediatric MRI study. *Genes, Brain Behav*. 2010;9:288–295. doi:10.1111/j.1601-183X.2009.00558.x.

21. Clark D, Royer H. The Effect of Education on Adult Mortality and Health: Evidence from Britain. *Am Econ Rev*. 2013;103(6):2087–2120. doi:10.1257/aer.103.6.2087.

22. Lleras-Muney A. The relationship between education and adult mortality in the United States. *Rev Econ Stud*. 2005;72:189–221. doi:10.1111/0034-6527.00329.

23. Angrist JD, Keueger AB. Does Compulsory School Attendance Affect Schooling and Earnings? *Q J Econ*. 1991;106:979–1014. doi:10.2307/2937954.

24. Conrad KM, Flay BR, Hill D. Why children start smoking cigarettes: predictors of onset. *Br J Addict*. 1992;87:1711–1724. doi:10.1111/j.1360-0443.1992.tb02684.x.

25. Chassin L, Presson CC, Rose JS, Sherman SJ. The natural history of cigarette smoking from adolescence to adulthood: demographic predictors of continuity and change. *Health Psychol*. 1996;15:478–484. doi:10.1037/0278-6133.15.6.478.

Table 1: Descriptive statistics of the MIDUS II and Minnesota twins samples stratified by zygocity

	Total		MZ Twins		DZ Twins	
	Mean	SD	Mean	SD	Mean	SD
A. MIDUS II twins						
Age	53.22	11.24	52.66	10.99	53.92	11.54
Male	0.39	--	0.44	--	0.33	--
Years of education	14.17	2.61	14.32	2.48	13.98	2.76
Body mass index	27.38	5.31	27.24	5.41	27.56	5.20
Cigs/day	5.05	11.07	5.43	11.59	4.57	10.38
Health visits/year	1.71	2.05	1.61	1.89	1.84	2.25
N	528		296		232	
B. Minnesota twins						
Age	47.09	5.62	46.67	5.51	47.62	5.71
Male	0.35	--	0.36	--	0.34	--
Years of education	13.69	2.65	13.88	2.70	13.45	2.56
Body mass index	25.84	4.65	25.78	4.55	25.91	4.78
N	2350		1318		1032	

Table 2: Parameter estimates for the effect of schooling, genetics, shared environments, and individual environments on health behaviors

	BMI		Cigarettes/day		Health visits/year	
	Coef	SE	Coef	SE	Coef	SE
A. MIDUS II twins						
Years of schooling	0.061	0.089	0.049	0.088	0.046	0.123
Genetic						
ayx	0.174	0.201	-0.175	0.162	0.271	0.231
ayy	0.715	0.048	0.000	0.251	0.376	0.179
axx	0.567	0.113	0.576	0.110	0.568	0.112
Shared environments						
cyx	-0.129	0.142	-0.106	0.126	0.062	0.131
cyy	0.000	0.482	0.563	0.046	0.000	0.436
cxx	0.603	0.101	0.597	0.102	0.604	0.100
Individual environments						
eyy	0.576	0.032	0.629	0.028	0.761	0.042
exx	0.533	0.032	0.530	0.031	0.532	0.031
B. Minnesota Twin Registry						
Years of schooling	0.018	0.037				
Genetic						
ayx	-0.113	0.078				
ayy	0.823	0.022				
axx	0.718	0.055				
Shared environments						
cyx	-0.091	0.108				
cyy	0.000	0.179				
cxx	0.362	0.101				
Individual environments						
eyy	0.564	0.015				
exx	0.579	0.016				

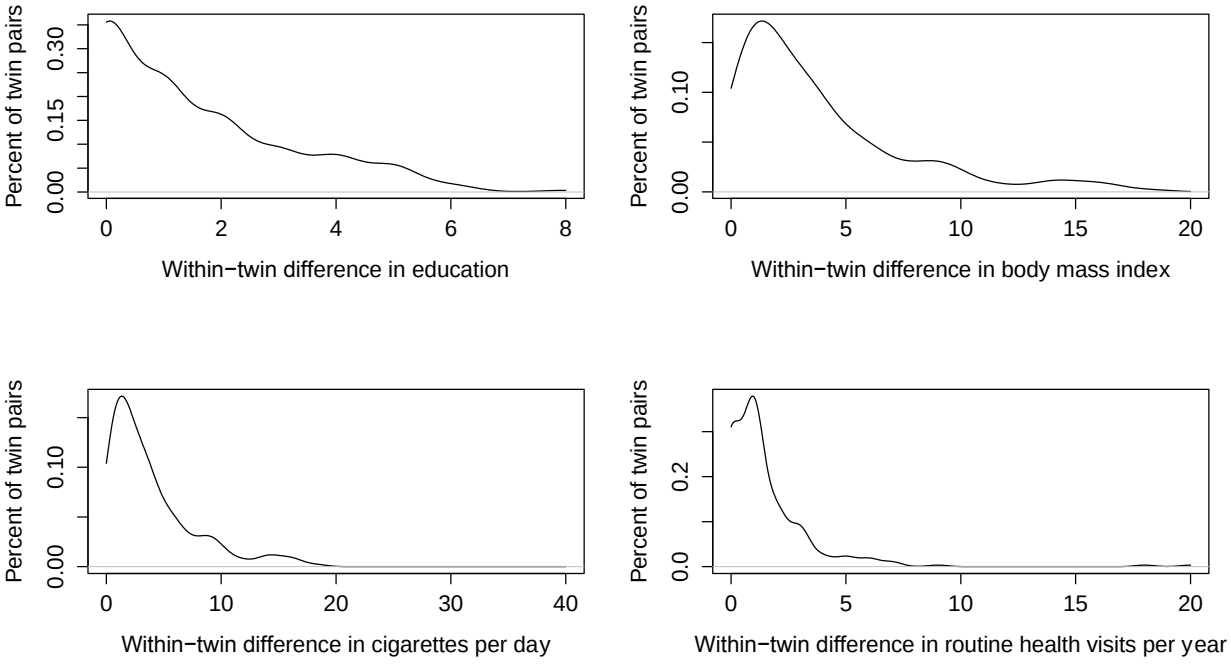
Notes: Parameters were estimated using an ACE-B structural equation model

Table 3: Decomposition of the total variance in health behaviors into genetics, shared environments, and individual environments

	BMI	Cigarettes/day	Health visits/year
A. MIDUS II twins			
Genetics	60.80%	2.90%	25.50%
Shared environments	1.00%	43.60%	1.00%
Individual environments	38.20%	53.50%	73.40%
B. Minnesota twins			
Genetics	67.90%		
Shared environments	0.70%		
Individual environments	31.40%		

Figure 1

A. Distribution of within-twin difference in schooling and health behaviors for MIDUSII twins, 262 twin pairs



B. Distribution of within-twin difference in schooling and body mass index for Minnesota twins, 1175 twin pairs