Using Marginal Structural Models to Estimate the Direct Effect of Adverse Childhood Social Conditions on Onset of Heart Disease, Diabetes, and Stroke

Arijit Nandi,[a] *M. Maria Glymour,*[b] *Ichiro Kawachi,*[b] *and Tyler J. VanderWeele*[c,d]

Background: Early-life socioeconomic status (SES) is associated with adult chronic disease, but it is unclear whether this effect is mediated entirely via adult SES or whether there is a direct effect of adverse early-life SES on adult disease. Major challenges in evaluating these alternatives include imprecise measurement of early-life SES and bias in conventional regression methods to assess mediation. In particular, conventional regression approaches to direct effect estimation are biased when there is time-varying confounding of the association between adult SES and chronic disease by chronic disease risk factors.

Methods: First-reported heart disease, diabetes, and stroke diagnoses were assessed in a national sample of 9760 Health and Retirement Study participants followed biennially from 1992 through 2006. Early-life and adult SES measures were derived using exploratory and confirmatory factor analysis. Early-life SES was measured by parental education, father's occupation, region of birth, and childhood rural residence. Adult SES was measured by respondent's education, occupation, labor force status, household income, and household wealth. Using marginal structural models, we estimated the direct effect of early-life SES on chronic disease onset that was not mediated by adult SES. Marginal structural models were estimated with stabilized inverse probability-weighted log-linear models to adjust for risk factors that may have confounded associations between adult SES and chronic disease.

Results: During follow-up, 24%, 18%, and 9% of participants experienced first onset of heart disease, diabetes, and stroke, respectively. Comparing those in the most disadvantaged with the least disadvantaged quartile, early-life SES was associated with coronary heart disease (risk ratio = 1.30 [95% confidence interval = 1.12–1.51]) and diabetes (1.23 [1.02–1.48]) and marginally associated with stroke via pathways not mediated by adult SES.

Conclusions: Our results suggest that early-life socioeconomic experiences directly influence adult chronic disease outcomes.

(*Epidemiology* 2012;23: 223–232)

Early-life socioeconomic conditions predict adult cardiovascular disease (CVD); however, the role of adult socioeconomic status (SES) in mediating this relationship remains controversial, and persistent socioeconomic inequalities in chronic diseases exist.[1–7] At least 3 alternative hypotheses linking early-life SES to chronic disease outcomes in later life have been proposed.[1] According to the latent-effects model, early-life SES affects risk for chronic disease independent of adult SES. The effect of early-life SES on chronic disease that is not mediated by adult SES (eg, if everyone in the population was fixed to a "high" level of adult SES) is defined as the controlled direct effect of early-life SES[8]; direct effects may include physiologic embodiment of socioeconomic disadvantage and attendant exposures through "fetal-programming" (represented by path 1 in the Figure, where path 2*3 represents the effect of early-life SES on chronic disease outcomes mediated by risk factors) and behavioral changes and shifts in chronic disease risk factors (represented by path 2*3 in the Figure).[9] In contrast, the pathway or social trajectories model suggests that early-life SES does not directly affect disease risk. Rather, early-life SES is linked to chronic disease in later life through its influence on SES in adulthood (represented by paths 4*5 and 2*6*5 in the Figure). Finally, the cumulative exposure model hypothesizes that SES in both early-life and adulthood directly influences disease risk in later life.

Extant research concerning the relation between life course SES and chronic disease is mixed. Some evidence suggests that early-life SES directly influences CVD independent of adult SES.[10–13] However, several studies have shown that associations between early-life SES and CVD outcomes were atten-

Submitted 15 March 2011; accepted 25 October 2011.

From the [a]Institute for Health and Social Policy and Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC Canada; and Departments of [b]Society, Human Development, and Health, [c]Epidemiology, and [d]Biostatistics, Harvard School of Public Health, Boston, MA.

Supported by the Robert Wood Johnson Health & Society Scholars Program (to A.N.); the American Heart Association (10SDG264024 to M.M.G.); and the National Institutes of Health (HD060696 to T.J.V.). The authors reported no other financial interests related to this research.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com). This content is not peer-reviewed or copy-edited; it is the sole responsibility of the author.

Editors' note: *A commentary on this article appears on page 233.*

Correspondence: Arijit Nandi, Institute for Health and Social Policy & Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, 1130 Pine Ave West, Montreal, QC H3A 1A3 Canada. E-mail: arijit.nandi@mcgill.ca.

ISSN: 1044-3983/12/2302-0223
DOI: 10.1097/EDE.0b013e31824570bd

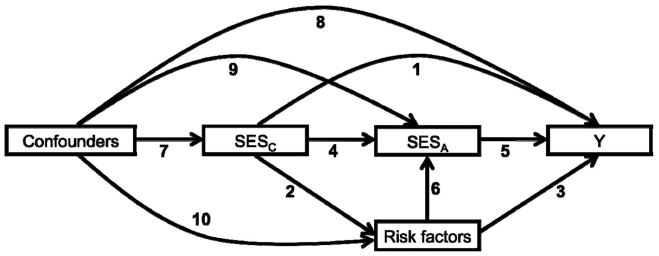


FIGURE. Causal diagram of the hypothesized effects of early-life (SES_C) and adult SES (SES_A) on chronic disease outcomes (Y) where "confounders" denote potential sociodemographic confounders (ie, age, race, sex, self-rated childhood health) of the association between early-life SES, adult SES and the chronic disease outcomes, and "risk factors" denote conventional chronic disease risk factors (ie, high blood pressure, BMI, self-rated health, diabetes, functional limitations, current smoking, alcohol consumption) that may operate as mediators of the association between early-life SES and health and also as confounders of the association between adult SES and health.

uated after adjustment for adult SES and CVD risk factors.[14,15] These inconsistencies may result from methodological challenges in estimating direct effects. For example, greater error in the measurement of early-life relative to adult SES may underestimate the direct effect of early-life SES on chronic disease outcomes.[16] Additionally, conventional regression methods to assess mediation,[17] specifically by regressing a chronic disease outcome on early-life SES, the proposed mediator (ie, adult SES), and potential confounders and chronic disease risk factors (eg, body mass index), produces an unbiased estimate of the controlled direct effect of early-life SES only under very specific conditions.[18]

In this study, we aimed to assess whether the effects of early-life SES on adult onset of coronary heart disease (CHD), diabetes, and stroke were mediated by adult SES using data from the US-based Health and Retirement Study. We improved on extant work by using factor analyses to measure SES and applying a marginal structural model to control for adult cardiovascular risk factors. Adult risk factors may confound the association of adult SES and CVD but mediate the effect of early-life SES on adult CVD. Conventional approaches to direct effect estimation are biased when there is time-varying confounding by adult risk factors. However, by handling potential confounding by measured covariates through weighting rather than conditioning on covariates, marginal structural models allow for the identification of direct effects even in settings in which conventional approaches are biased, including when there is a consequence of early-life SES that confounds the association between adult SES and the outcome of interest.[19,20] We compare results from marginal structural models with estimates obtained using the conventional direct effects estimation procedure based on regression adjustment for measured risk factors.[17]

METHODS

Study Design

The Health and Retirement Study is a longitudinal, biennial survey of a national sample of US adults born in 1931–1941 and their spouses. The study was initiated in 1992, based on a multistage area probability sample. The initial response rate was 81%, and biennial follow-up interviews (or proxy for decedents) were conducted between the baseline assessments in 1992 and 2006, with wave-to-wave retention rates through 2006 of approximately 90%.[21] The University of Michigan Health Sciences Human Subjects Committee approved the Health and Retirement Study.

We followed respondents through 2006 for first onset of 3 chronic disease outcomes: (1) fatal or nonfatal heart disease; (2) diabetes; or (3) fatal or nonfatal stroke. The Health and Retirement Study most commonly learns of the death of a respondent when an interviewer attempts to reach the respondent for a biennial follow-up interview. In these instances, the respondent's spouse or another close family member completes a final (exit) interview. The overall response rate for exit interviews conducted through 2002 was 93%, and there is no evidence of systematic bias by demographic groups. The Health and Retirement Study conducts linkages using the Social Security Death Index and the National Death Index to confirm the status, timing, and cause of death for deceased respondents. Further details of the study design and outcomes assessment of the Health and Retirement Study are available elsewhere.[22–24]

Among 9760 age-eligible Health and Retirement Study participants enrolled in 1992, we excluded 495 (5%) without a follow-up interview between 1994 and 2006. Excluding prevalent cases of each outcome at baseline, we defined 3 separate samples, comprising 8318 CHD-free, 8278 diabetes-free, and 9055 stroke-free participants.

Measures

For consistency with previous studies, we used standard RAND coding of wealth, income, and other variables, when available.[25]

Chronic Disease Outcomes

At each wave, participants were asked, "Since we last talked to you, has a doctor told you that you had a heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems?" Diabetes and stroke were similarly assessed. For deceased participants or those unavailable for direct interviews, proxy informants, typically spouses, were interviewed.

Life-course Socioeconomic Status

We used exploratory and confirmatory factor analyses to estimate the latent construct SES at 2 stages in the life-course, early life and adulthood. We selected 8 indicators of early-life SES based on retrospective

assessment and 6 indicators of adult SES (eTable 1, http://links.lww.com/EDE/A552).[26,27] Exploratory factor analysis with quartimin rotation was used to assess the dimensionality of our SES indicators, resulting in a final two-factor model. Indicators of early-life SES included father's educational attainment (0–7, 8–11, 12, >12 years), mother's educational attainment (0–7, 8–11, 12, >12 years), father's occupation (manual/unskilled service, professional/white collar), birth in the southern United States (yes/no), and rural residence during childhood (yes/no); all were assessed retrospectively at baseline with the exception of father's occupation, which was assessed in 1998. Indicators of adult SES included respondent's educational attainment (<high school, high school/GED, some college, college+), respondent's longest job occupation (manual/unskilled service, professional/white collar), respondent's current labor force status (works full-time/part-time/retired, unemployed/disabled/not in labor force), household income (split into quartiles), and household wealth (split into quartiles), all assessed at baseline. We next used confirmatory factor analysis with maximum likelihood estimation to output factor scores for early-life and adult SES. Continuous early-life and adult SES scores were cut into quartiles for analysis. Factor analyses were conducted in Mplus version 5. Further detail concerning the methods used to measure life course SES is provided in the eAppendix, http://links.lww.com/EDE/A552.

Confounders and Risk Factors

We distinguish between (1) sociodemographic confounders that potentially confound the relation between early-life SES and our outcomes and (2) chronic disease risk factors that potentially confound the relation between adult SES and our outcomes (as shown in the Figure).[28,29] Measured confounders of the relation between early-life SES and our outcomes included baseline age (categorized as 50–52, 53–55, 56–58, or >58), sex, race (dichotomized as black or other), and self-rated childhood health (very good/excellent, good, fair/poor, or unknown/missing). Measured baseline risk factors that potentially confounded the relation between adult SES and our chronic disease outcomes included whether a doctor had ever told the respondent that they had high blood pressure (no or yes), body mass index (BMI) (categorized as <25, 25–29.9, or 30+), self-rated health (categorized as excellent, very good, good, fair, or poor), diabetes (no or yes), whether health problems limited the respondent's ability to work (no or yes), current smoking (no or yes), and alcohol consumption (categorized as doesn't drink, <1 drink/day, 1–2 drinks/day, or ≥3 drinks/day). Diabetes was not included as a risk factor in analyses for which diabetes was the outcome.

Statistical Analyses

Our goal was to estimate, for individual i, the controlled direct effect of early-life SES (A_i) on chronic disease outcomes (Y_i) that was not mediated by adult SES (M_i) after accounting for potential confounding by measured confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health) and risk factors (ie, high blood pressure, BMI, self-rated health, diabetes, functional limitations, current smoking, alcohol consumption), represented by the vectors C_i and R_i, respectively. Here, and throughout, the variables A_i and M_i denote the vectors of indicator functions for the various quartiles of early-life and adult SES, respectively.

Formally, if we allow $Y_i(a,m)$ to represent the counterfactual outcome Y_i if potentially counter to the fact A_i had been set to a and M_i had been set to m, the controlled direct effect on the risk ratio (RR) scale comparing $A_i = a$ with $A_i = a^*$ for some fixed level $M_i = m$ is defined as

$$RR_i^m(a,a^*) = \frac{P[Y_i(a,m) = 1]}{P[Y_i(a^*,m) = 1].}$$

Although it is often assumed that the effect of A_i on Y_i is homogenous for all levels of M_i, the controlled direct effect of early-life SES may depend on the value to which adult SES is set. As such, we tested for interaction between A_i and M_i on Y_i in models estimating direct effects of early-life SES. However, in the absence of such interaction, the controlled direct effect is defined as the effect of early-life SES on the particular chronic disease outcome if adult SES were set to the highest quartile; this effect includes both paths 1 and 2*3 in the Figure, but not paths via adult SES (paths 4*5 or 2*6*5).

We compare 2 approaches for estimating the direct effect of early-life SES on our outcomes: the conventional regression approach to effect decomposition[17] and the marginal structural model approach using inverse probability weights.[19,20] First, we estimated on the RR scale the total effect of early-life SES on chronic disease outcomes, conditional on $C_i = c$,

$$RR_i(a,a^*|c) = \frac{P[Y_i(a) = 1|C_i = c]}{P[Y_i(a^*) = 1|C_i = c],}$$

by fitting a log-linear model for each outcome ($Y_i = 1$ if participant reports the outcome during follow-up, 1992–2006) regressed on quartiles of early-life SES (A_i) and the set of confounders (C_i), separately for the 3 samples comprising CHD-free, diabetes-free, and stroke-free respondents:

$$\log[P(Y_i = 1|A_i = a, C_i = c)] = \beta_0 + \beta_1' A_i + \beta_2' C_i. \quad (1)$$

The coefficient β_1 in Model 1 gives an estimate of the total effect of early-life SES provided that the measured confounders suffice to control for confounding between early-life SES and the chronic disease outcomes. Second, we estimated the direct effect of early-life SES on

our outcomes using the approach proposed by Baron and Kenny (1986).[17] Specifically, we regressed each outcome Y_i on quartiles of early-life SES (A_i), quartiles of adult SES (M_i), and the set of sociodemographic confounders (C_i), by fitting a log-linear model of the form:

$$\log[P(Y_i = 1|A_i = a, M_i = m, C_i = c)] = \beta_0 + \beta_1' A_i + \beta_2' M_i + \beta_3' C_i. \quad (2)$$

The coefficient β_1 in Model 2 gives an estimate of the direct effect of early-life SES not through pathways involving adult SES (paths 1 and 2*3 in the Figure) provided that the measured confounders suffice to control for confounding between (i) early-life SES and the chronic disease outcomes and (ii) adult SES and the chronic disease outcomes. There was no evidence of interaction between C_i and A_i. As such, the coefficient β_1 gives both conditional direct effects and also the marginal direct effect described earlier. We additionally tested for interaction between early-life and adult SES on the chronic disease outcomes by including cross-product terms between A_i and M_i in Model 2. As observed in eTable 2 (http://links.lww.com/EDE/A552), there was no evidence of interaction beyond what would be expected by chance alone for any of our outcomes, and the cross-product terms were subsequently omitted.

To compare with prior work in which adult risk factors were controlled, we also fit a log-linear regression model (Model 3) in which adult risk factors (R_i) were included as covariates:

$$\log[P(Y_i = 1|A_i = a, M_i = m, C_i = c, R_i = r)] = \beta_0 + \beta_1' A_i + \beta_2' M_i + \beta_3' C_i + \beta_4' R_i. \quad (3)$$

The coefficient β_1 in Model 3 gives an estimate of the direct effect of early-life SES not through pathways involving adult SES or the adult risk factors (path 1 in the Figure) provided that measured confounders and risk factors suffice to control for confounding between (i) early-life SES and the chronic disease outcomes, (ii) adult SES and the chronic disease outcomes, and (iii) the adult risk factors and chronic disease outcomes. Note that even under these assumptions, this is a different direct effect from that defined earlier (ie, it includes only the effect of path 1 in the Figure). There was no consistent evidence of interaction between A_i and M_i in Model 3 (eTable 3, http://links.lww.com/EDE/A552) for any of our outcomes, and therefore we omitted the cross-product terms.

The controlled direct effect of early-life SES on our outcomes using a marginal structural model (Model 4) was estimated by a stabilized inverse-probability-weighted marginal structural model as described by VanderWeele and Robins et al.[19,20] In brief, we fit a weighted log-linear regression model of the form:

$$\log[P(Y_i = 1|A_i = a, M_i = m)] = \beta_0 + \beta_1' A_i + \beta_2' M_i, \quad (4)$$

excluding a cross-product term between early-life and adult SES because there was no consistent evidence of interaction for any of the outcomes (eTable 4, http://links.lww.com/EDE/A552). Potential confounding by the sets of confounders and risk factors was accounted for by fitting the earlier model with stabilized inverse probability weights of the form $W = w_i^A * w_i^M$, where

$$w_i^A = \frac{P(A = a_i)}{P(A = a_i|C = c_i)} \quad (5)$$

and

$$w_i^M = \frac{P(M = m_i|A = a_i)}{P(M = m_i|A = a_i, C = c_i, R = r_i)}. \quad (6)$$

The weight w_i^A accounts for measured confounding of the relation between early-life SES and the chronic disease outcomes, and the weight w_i^M accounts for measured confounding of the relation between adult SES and the outcomes. The coefficients β_1 in the final weighted model (Model 4) give the direct effect of early-life SES not through adult SES pathways provided that (i) the measured confounders suffice to control for confounding between early-life SES and chronic disease outcomes and (ii) the measured confounders, early-life SES, and the adult risk factors suffice to control for confounding between adult SES and chronic disease outcomes. Note that under these assumptions, the direct effect β_1 includes the effect of early-life SES through adult risk factor pathways (path 2*3 in the Figure, in addition to path 1), but not through adult SES.

The denominator of w_i^A is the probability of receiving the quartile of early-life SES the individual in fact experienced, conditional on the set of confounders. The denominator of w_i^M is the probability of receiving the quartile of the mediator, adult SES, the individual in fact experienced, conditional on quartiles of early-life SES, confounders, and the set of risk factors. Stabilizing the weights by including probabilities in the numerator results in more efficient estimation.[20] Weights were estimated using multinomial logistic regressions where the response variables for w_i^A and w_i^M were quartiles of SES in early-life and adulthood, respectively. Predicted probabilities for the numerator and denominator were assigned based on the category of actual SES reported and divided to obtain stabilized weights. The distributions of the multiplied weights, W, were as follows: (1) mean =1.00, range = 0.15–10.88, IQR = 0.68–1.12 for heart disease; (2) mean = 1.00, range = 0.16–9.88, IQR = 0.69–1.13 for diabetes; and (3) mean = 1.00, range = 0.16–10.12, IQR = 0.68–1.12 for stroke. See eFigures

2–4 (http://links.lww.com/EDE/A552) for histograms and additional statistics describing the distributions of the stabilized weights for each of the outcomes. All analyses were conducted using SAS version 9.1.3. Annotated SAS code for estimating the models described earlier is provided in the eAppendix (http://links.lww.com/EDE/A552).

RESULTS

Characteristics of the total sample and the samples used in analyses for each condition are shown in Table 1. Between baseline and the end of follow-up in 2006, 24% of participants experienced first-onset heart disease (1886 nonfatal and 115 fatal events), 18% experienced first-onset diabetes (1461 events), and 8% experienced first-onset stroke (618 nonfatal and 90 fatal events).

Estimates of the total and direct effects of early-life SES on CHD are shown in Table 2. The total effects model (Model 1) indicated that those in the third and fourth most disadvantaged quartiles of early-life SES had, respectively, 12% (RR = 1.12 [95% confidence interval (CI) = 1.00–1.25]) and 16% (1.16 [1.04–1.30]) increased risk of CHD compared with those in the least disadvantaged quartile of early-life SES. After adjusting for adult SES in Model 2 as in a conventional regression approach, there was a null association between early-life SES and CHD (1.00 [0.87–1.16]), comparing those in the most disadvantaged to least disadvantaged quartile of early-life SES. Results were essentially unchanged after adjusting for adult risk factors (Model 3). Applying the marginal structural model approach using inverse probability weights to adjust for adult risk factors (Model 4), those in the third and fourth most disadvantaged quartiles of early-life SES were estimated to have, respectively, 23% (1.23 [1.08–1.40]) and 30% (1.30 [1.12–1.51]) increased risk of CHD compared with those in the least disadvantaged quartile of early-life SES.

Estimates of the total and direct effects of early-life SES on diabetes are shown in Table 3. There was a dose-response association between early-life SES and diabetes (Model 1). Compared with those in the least disadvantaged quartile of early-life SES, those in the second, third, and fourth quartiles had, respectively, 12% (1.12 [0.97–1.29]), 31% (1.31 [1.14–1.51]), and 62% (1.62 [1.41–1.86]) increased risk of diabetes. The association between early-life SES and diabetes was substantially attenuated after adjustment for adult SES (Model 2) and nearly null after additional adjustment for adult risk factors (Models 3). In the marginal structural model (Model 4), those who experienced the greatest disadvantage in early life showed increased risk of diabetes relative to those in the least disadvantaged quartile (1.23 [1.02–1.48]).

Estimates of the total and direct effects of early-life SES on stroke are shown in Table 4. The total-effects model (Model 1) showed that those in the third and fourth most disadvantaged quartiles of early-life SES had, respectively, 37% (1.37 [1.11–1.69]) and 59% (1.59 [1.29–1.95]) increased risk of stroke relative to those in the least disadvantaged quartile. After controlling for adult SES and risk factors using the conventional regression approach (Models 2 and 3), associations between early-life SES and stroke were attenuated and provided no evidence of a direct effect of early-life SES. The associations between early-life SES and stroke using the marginal structural model approach (Model 4) were stronger in magnitude than for the conventional regression approach; for example, the fourth quartile of early-life SES was associated with an RR for stroke of 1.29 (95% CI = 0.96–1.72) compared with those in the top quartile.

DISCUSSION

Consistent with several prior studies, we found evidence of a total effect of early-life SES on incidence of CHD, diabetes, and stroke in a sample of 9760 Health and Retirement Study enrollees.[30–32] In a novel application of marginal structural models, we also found evidence of pathways from early-life SES to each outcome not mediated by adult SES.

Estimating the effect of early-life SES on risk of adult chronic disease via mechanisms other than adult SES raises several methodological issues primarily because early-life SES strongly influences adult SES. When we used the traditional direct effects estimation approach[17] of regressing chronic disease outcomes on early-life SES and adult SES, we found that estimated effects of early-life SES on our outcomes were attenuated, and there was no evidence of a direct effect. The pattern was similar whether or not we additionally adjusted for adult CVD risk factors. However, there are several adult risk factors that may mediate the effect of early-life SES but confound the association of adult SES with our health outcomes (as shown in the Figure). For example, prior work suggests that chronic disease risk factors, including obesity, may lie on the pathway between early-life SES and adult health and also be causally related to adult SES, suggesting that these risk factors may operate contemporaneously as mediators of the association between early-life SES and health, as well as confounders of the association between adult SES and health.[33–35]

When there is a consequence of early-life SES that confounds the association between adult SES and the outcome of interest (if paths 2, 3, and 6 in the Figure exist), controlled direct effects cannot be identified using conventional regression approaches, regardless of whether the model is additionally adjusted for adult risk factors.[19] When factors that confound the association between adult SES and chronic disease outcomes are omitted from the regression model, then conditioning on adult SES may induce bias in the estimate of the direct effect of early-life SES due to collider stratification.[36] This bias may have resulted in an underestimation of the direct effect of early-life SES in our analyses, although

TABLE 1. Descriptive Characteristics and Bivariate Associations Between Sociodemographic Characteristics, Risk Factors, Life-course SES and Chronic Disease (ie, Heart Disease, Diabetes, and Stroke) Outcomes; n = 9760 Health and Retirement Study Enrollees Followed From 1992 to 2006

	Total No. (% of Total)	Heart Disease Sample No. (% With CHD)[a]	Diabetes Sample No. (% With Diabetes)[a]	Stroke Sample No. (% With Stroke)[a]
Total sample size	9760	8318 (24)[b]	8278 (18)[b]	9055 (8)[b]
Sociodemographic confounders				
Birth cohort/baseline age (years)				
50–52.9	2175 (22)	1910 (21)	1891 (17)	2048 (6)
53–55.9	2757 (28)	2387 (22)	2394 (17)	2560 (6)
56–58.9	2606 (27)	2221 (25)	2188 (19)	2433 (8)
59+	2222 (23)	1800 (29)	1805 (17)	2014 (10)
Race/ethnicity				
White/other	8065 (83)	6906 (24)	6968 (16)	7528 (7)
Black	1695 (17)	1412 (24)	1310 (25)	1527 (12)
Sex				
Male	4594 (47)	3796 (27)	3868 (18)	4210 (9)
Female	5166 (53)	4522 (21)	4410 (17)	4845 (7)
Childhood self-rated health				
Very good/excellent	5861 (60)	5374 (24)	5314 (18)	5747 (7)
Good	1478 (15)	1315 (28)	1307 (22)	1445 (10)
Fair/poor	548 (6)	28.91 (29)	483 (23)	526 (11)
Unknown/missing	1873 (19)	17.28 (17)	1174 (11)	1337 (7)
Risk factors				
High blood pressure				
No	5884 (60)	5247 (20)	5271 (13)	5548 (5)
Yes	3876 (40)	3071 (32)	3007 (25)	3507 (12)
Body mass index (BMI; kg/m^2)				
<25	3419 (35)	2980 (19)	3073 (7)	3181 (7)
25–29.9	4003 (41)	3417 (24)	3415 (18)	3712 (8)
30+	2338 (24)	1921 (31)	1790 (36)	2162 (10)
Self-rated health				
Poor	2119 (22)	2002 (15)	1999 (11)	2037 (4)
Fair	2723 (28)	2506 (22)	2490 (15)	2601 (5)
Good	2711 (28)	2313 (26)	2291 (21)	2535 (9)
Very good	1405 (14)	1031 (36)	1009 (27)	1255 (13)
Excellent	802 (8)	466 (38)	489 (27)	626 (17)
Diabetes				
No	8696 (89)	7542 (22)		8124 (7)
Yes	1064 (11)	776 (41)		931 (17)
Health problems limit work				
No	7645 (79)	6875 (22)	6750 (16)	7259 (6)
Yes	2092 (21)	1433 (35)	1517 (23)	1782 (14)
Current smoker				
No	7084 (73)	6077 (22)	6000 (17)	6602 (7)
Yes	2676 (27)	2241 (28)	2278 (19)	2453 (11)
Alcohol use (drinks/day)				
0	3871 (40)	3210 (26)	3077 (21)	3520 (10)
<1	4346 (45)	3779 (23)	3833 (16)	4101 (6)
1–2	1024 (10)	894 (21)	909 (13)	957 (7)
≥3	520 (5)	435 (25)	459 (14)	477 (10)
Socioeconomic status				
Early-life SES[b]				
Quartile 1 (highest SES)	2443 (25)	2160 (22)	2167 (14)	2298 (6)
Quartile 2	2437 (25)	2104 (23)	2129 (15)	2279 (5)
Quartile 3	2440 (25)	2038 (25)	2063 (18)	2263 (9)
Quartile 4 (lowest SES)	2440 (25)	2016 (26)	1919 (24)	2215 (11)
Continuous factor score mean (SD)	0.00 (2.71)	0.08 (2.72)	0.13 (2.69)	0.04 (2.71)

(*Continued*)

TABLE 1. (*Continued*)

	Total No. (% of Total)	Heart Disease Sample No. (% With CHD)[a]	Diabetes Sample No. (% With Diabetes)[a]	Stroke Sample No. (% With Stroke)[a]
Adult SES[b]				
Quartile 1 (highest SES)	2444 (25)	2183 (22)	2198 (13)	2321 (5)
Quartile 2	2440 (25)	2130 (24)	2147 (15)	2294 (6)
Quartile 3	2436 (25)	2020 (23)	2040 (18)	2245 (8)
Quartile 4 (lowest SES)	2440 (25)	1985 (28)	1893 (26)	2195 (12)
Continuous factor score mean (SD)	0.00 (1.43)	0.06 (1.43)	0.09 (1.41)	0.03 (1.43)

[a]No. of participants without each condition in 1992; numbers in parentheses show percent experiencing first onset of condition between 1992 and 2006.
[b]Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles of early-life and adult SES.
SD indicates standard deviation.

TABLE 2. Total Effects, Conventional Regression With and Without Adjustment for Risk Factors, and MSM Results for Heart Disease (n = 8308)

Life-course Socioeconomic Status	Total Effects (Model 1)[a] RR (95% CI)	Regression Adjustment for Adult SES and Confounders (Model 2)[b] RR (95% CI)	Regression Adjustment for Adult SES, Confounders, and Risk Factors (Model 3)[c] RR (95% CI)	Marginal Structural Model (Model 4)[d] RR (95% CI)
Early-life SES[e]				
Quartile 1 (highest SES)[f]	1.00	1.00	1.00	1.00
Quartile 2	1.01 (0.90–1.13)	0.98 (0.87–1.10)	1.00 (0.87–1.14)	1.05 (0.93–1.18)
Quartile 3	1.12 (1.00–1.25)	1.04 (0.91–1.19)	1.05 (0.91–1.22)	1.23 (1.08–1.40)
Quartile 4 (lowest SES)	1.16 (1.04–1.30)	1.00 (0.87–1.16)	1.02 (0.86–1.21)	1.30 (1.12–1.51)
Adult SES[a]				
Quartile 1 (highest SES)[f]		1.00	1.00	1.00
Quartile 2		1.09 (0.97–1.22)	0.95 (0.83–1.09)	0.98 (0.87–1.09)
Quartile 3		1.04 (0.91–1.19)	0.83 (0.71–0.97)	0.85 (0.75–0.97)
Quartile 4 (lowest SES)		1.29 (1.11–1.49)	0.87 (0.73–1.05)	0.86 (0.74–1.00)

[a]Adjusted for early-life SES and set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health).
[b]Adjusted for early-life SES, adult SES, and set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health).
[c]Adjusted for early-life SES, adult SES, set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health), and set of risk factors (ie, high blood pressure, BMI, self-rated health, diabetes, functional limitations, current smoking, alcohol consumption).
[d]Stabilized inverse probability weights used to account for potential confounding by the sets of confounders and risk factors.
[e]Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles of early-life and adult SES.
[f]Reference category.

collider-stratification bias could be in either direction. However, adjusting for such chronic disease risk factors in a conventional regression approach blocks measured risk factor pathways linking early-life SES to chronic disease. This would estimate the effect of early-life SES on chronic disease that is not through adult SES and not through the adult risk factors. The marginal structural model circumvents this confounding, and our results suggest that conflicting results in the existing literature might be explained, in part, by different biases that result with and without controlling for adult risk factors using conventional regression approaches.

There are a number of caveats to our study. First, there may be error in the measurement of self-reported SES or health outcomes. Second, we used full information maximum likelihood to obtain factors scores for early-life SES despite the fact that some indicators were missing. This method has been shown to be unbiased and more efficient than other common methods for handling missing data in structural equation modeling.[37] Third, we assume that our parametric regression and marginal structural models were correctly specified and that the consistency assumption holds.[38] Future work should consider alternative estimation approaches for direct effects that rely on weaker modeling assumptions.[39–41] Fourth, marginal structural models and inverse-probability-of-treatment weighting require a positivity assumption that the probabilities in the denominator of the weights are nonzero,[42] and this assumption may be violated if there were negligible levels of social mobility in the Health and Retirement Study sample. However, alternative categorizations of SES did not substantively change our findings. Fifth, use of inverse probability weights does not address unmeasured confounding. We attempted to minimize unmeasured con-

TABLE 3. Total Effects, Conventional Regression With and Without Adjustment for Risk Factors, and MSM Results for Diabetes (n = 8267)

Life-course Socioeconomic Status	Total Effects (Model 1)[a] RR (95% CI)	Regression Adjustment for Adult SES and Confounders (Model 2)[b] RR (95% CI)	Regression Adjustment for Adult SES, Confounders, and Risk Factors (Model 3)[c] RR (95% CI)	Marginal Structural Model (Model 4)[d] RR (95% CI)
Early-life SES[e]				
Quartile 1 (highest SES)[f]	1.00	1.00	1.00	1.00
Quartile 2	1.12 (0.97–1.29)	1.02 (0.87–1.19)	0.99 (0.83–1.17)	1.06 (0.91–1.23)
Quartile 3	1.31 (1.14–1.51)	1.07 (0.90–1.27)	0.99 (0.82–1.19)	1.15 (0.97–1.36)
Quartile 4 (lowest SES)	1.62 (1.41–1.86)	1.17 (0.98–1.41)	1.08 (0.88–1.32)	1.23 (1.02–1.48)
Adult SES[a]				
Quartile 1 (highest SES)[f]		1.00	1.00	1.00
Quartile 2		1.12 (0.95–1.31)	0.98 (0.83–1.17)	1.02 (0.87–1.18)
Quartile 3		1.27 (1.08–1.51)	1.06 (0.88–1.28)	1.02 (0.86–1.21)
Quartile 4 (lowest SES)		1.64 (1.37–1.98)	1.24 (1.00–1.54)	1.40 (1.17–1.68)

[a]Adjusted for early-life SES and set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health).
[b]Adjusted for early-life SES, adult SES, and set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health).
[c]Adjusted for early-life SES, adult SES, set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health), and set of risk factors (ie, high blood pressure, BMI, self-rated health, functional limitations, current smoking, alcohol consumption).
[d]Stabilized inverse probability weights used to account for potential confounding by the sets of confounders and risk factors.
[e]Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles of early-life and adult SES.
[f]Reference category.

TABLE 4. Total Effects, Conventional Regression With and Without Adjustment for Risk Factors, and MSM Results for Stroke (n = 9041)

Life-course Socioeconomic Status	Total Effects (Model 1)[a] RR (95% CI)	Regression Adjustment for Adult SES and Confounders (Model 2)[b] RR (95% CI)	Regression Adjustment for Adult SES, Confounders and Risk Factors (Model 3)[c] RR (95% CI)	Marginal Structural Model (Model 4)[d] RR (95% CI)
Early-life SES[e]				
Quartile 1 (highest SES)[f]	1.00	1.00	1.00	1.00
Quartile 2	0.86 (0.68–1.09)	0.71 (0.55–0.91)	0.74 (0.57–0.96)	0.75 (0.57–0.97)
Quartile 3	1.37 (1.11–1.69)	0.91 (0.70–1.18)	0.93 (0.71–1.21)	1.24 (0.95–1.61)
Quartile 4 (lowest SES)	1.59 (1.29–1.95)	0.88 (0.66–1.16)	0.90 (0.67–1.21)	1.29 (0.96–1.72)
Adult SES[a]				
Quartile 1 (highest SES)[f]		1.00	1.00	1.00
Quartile 2		1.17 (0.91–1.52)	0.98 (0.75–1.28)	0.98 (0.76–1.27)
Quartile 3		1.74 (1.33–2.28)	1.26 (0.95–1.67)	1.28 (0.98–1.68)
Quartile 4 (lowest SES)		2.35 (1.75–3.16)	1.35 (0.98–1.87)	1.49 (1.11–2.00)

[a]Adjusted for early-life SES and set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health).
[b]Adjusted for early-life SES, adult SES, and set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health).
[c]Adjusted for early-life SES, adult SES, set of confounders (ie, birth cohort, sex, race/ethnicity, and self-rated childhood health), and set of risk factors (ie, high blood pressure, BMI, self-rated health, diabetes, functional limitations, current smoking, alcohol consumption).
[d]Stabilized inverse probability weights used to account for potential confounding by the sets of confounders and risk factors.
[e]Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles of early-life and adult SES.
[f]Reference category.

founding in our study by including measures on recognized potential confounders included in the Health and Retirement Study. However, unmeasured risk factors, including mood disorders, may be caused by early-life SES and potentially confound the association between adult SES and our health outcomes. Alternatively, measured risk factors may be measured with error (eg, current smoking is an imperfect proxy for smoking history). Sixth, to be included in our sample, respondents must survive at least 50 years since birth (to be eligible for the Health and Retirement Study) and not be a prevalent case at baseline; individuals with lower SES are less likely to satisfy these criteria, and the exclusion of these individuals may underestimate the true effect of adult SES on health. Seventh, risk factors and adult SES were measured

concurrently. Although we included only those risk factors posited to influence adult SES, there may be a reciprocal relation between these characteristics, and further work should model this association longitudinally.

These caveats considered, our results support the hypothesis that early environmental conditions may have lasting effects on adult health and well-being. However, socioeconomic disadvantage during early life is a broad construct, and a number of mechanisms (encompassed by path 1 in the Figure) may "directly" link early-life SES to adult onset of chronic disease independent of intermediate levels of adult SES and measured risk factors. At the individual level, socioeconomic disadvantage is associated with neuroendocrine and inflammatory processes that may increase risk for chronic disease onset.[43,44] Although most of these studies have been conducted among adults, recent work suggests that socioeconomic disparities in subclinical risk for CVD may be established early in life[45]; whether these risks are explained by behavioral, physiologic, psychologic differences or some combination of these mechanisms, is unclear. Extending this to a multilevel framework that incorporates family, school, and neighborhood exposures, greater socioeconomic disadvantage in early life may engender differences in social, physical, or service environments and influence risk for chronic disease directly or indirectly.[46] The application of marginal structural models to longitudinal data and the extension of socioeconomic disadvantage to include place-based factors may help to elucidate the specific mechanisms linking the early-life socioeconomic environment to adult onset of chronic diseases.[47]

In conclusion, our analyses provide evidence for a direct effect of early-life SES on chronic disease in the Health and Retirement Study when using marginal structural models but not when using the conventional regression approach. Most theories of the life-course development of chronic disease suggest an underlying causal structure in which the marginal structural model would identify the causal effect of substantive interest, but conventional regression methods for estimating direct effects would not. These results therefore suggest that alternative methods, including marginal structural models, should be explored for estimating direct effects in the context of life-course epidemiology.

REFERENCES

1. Lynch J, Smith GD. A life course approach to chronic disease epidemiology. *Annu Rev Public Health*. 2005;26:1–35.
2. Fox CS, Evans JC, Larson MG, Kannel WB, Levy D. Temporal trends in coronary heart disease mortality and sudden cardiac death from 1950 to 1999: the Framingham Heart Study. *Circulation*. 2004;110:522–527.
3. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol*. 2002;31:285–293.
4. Braveman PA, Cubbin C, Egerter S, Williams DR, Pamuk E. Socioeconomic disparities in health in the United States: what the patterns tell us. *Am J Public Health*. 2010;100suppl 1:S186–S196.
5. Lloyd-Jones D, Adams R, Carnethon M, et al. Heart disease and stroke statistics–2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2009;119:e21–e181.
6. Mensah GA, Mokdad AH, Ford ES, Greenlund KJ, Croft JB. State of disparities in cardiovascular health in the United States. *Circulation*. 2005;111:1233–1241.
7. Saydah S, Lochner K. Socioeconomic status and risk of diabetes-related mortality in the U.S. *Public Health Rep*. 2010;125:377–388.
8. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992;3:143–155.
9. Kuh D, Davey Smith G. When is mortality risk determined? Historical insights into a current debate. *Soc Hist Med.* 1993;6:101–123.
10. Claussen B, Davey Smith G, Thelle D. Impact of childhood and adulthood socioeconomic position on cause specific mortality: the Oslo Mortality Study. *J Epidemiol Community Health*. 2003;57:40–45.
11. Smith GD, Hart C, Blane D, Hole D. Adverse socioeconomic conditions in childhood and cause specific adult mortality: prospective observational study. *BMJ*. 1998;316:1631–1635.
12. Smith GD, McCarron P, Okasha M, McEwen J. Social circumstances in childhood and cardiovascular disease mortality: prospective observational study of Glasgow University students. *J Epidemiol Community Health*. 2001;55:340–341.
13. Gliksman MD, Kawachi I, Hunter D, et al. Childhood socioeconomic status and risk of cardiovascular disease in middle aged US women: a prospective study. *J Epidemiol Community Health*. 1995;49:10–15.
14. Loucks EB, Lynch JW, Pilote L, et al. Life-course socioeconomic position and incidence of coronary heart disease: the Framingham Offspring Study. *Am J Epidemiol*. 2009;169:829–836.
15. Kaplan GA, Salonen JT. Socioeconomic conditions in childhood and ischaemic heart disease during middle age. *BMJ*. 1990;301:1121–1123.
16. Phillips AN, Smith GD. Bias in relative odds estimation owing to imprecise measurement of correlated exposures. *Stat Med*. 1992;11:953–961.
17. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51:1173–1182.
18. Kaufman JS, Maclehose RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiol Perspect Innov*. 2004;1:4.
19. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 2009;20:18–26.
20. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550–560.
21. Health and Retirement Study. Sample Sizes and Response Rates. Available at: http://hrsonline.isr.umich.edu/sitedocs/sampleresponse.pdf.
22. Juster FT, Suzman R. An overview of the Health and Retirement Study. *J Hum Resour*. 1995;(suppl 30):S7–S56.
23. Heeringa SG, Connor J. Technical description of the Health and Retirement Study sample design (Report No.: DR-002). Ann Arbor, MI: Survey Research Center, University of Michigan; 1995.
24. Ofstedal MB, Fisher GF, Herzog AR. *Documentation of Cognitive Functioning Measures in the Health and Retirement Study*. Ann Arbor, MI: Survey Research Center, University of Michigan; 2005.
25. St. Clair P, Blake D, Bugliari D, et al. RAND HRS Data Documentation, Version G. Rand Center for Study of Aging. Available at http://hrsonline.isr.umich.edu/index.php?p=showcbk.
26. Braveman PA, Cubbin C, Egerter S, et al. Socioeconomic status in health research: one size does not fit all. *JAMA*. 2005;294:2879–2888.
27. Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc*. 2007;99:1013–1023.
28. Pollitt RA, Rose KM, Kaufman JS. Evaluating the evidence for models of life course socioeconomic factors and cardiovascular outcomes: a systematic review. *BMC Public Health.* 2005;5:7.
29. Galobardes B, Smith GD, Lynch JW. Systematic review of the influence of childhood socioeconomic circumstances on risk for cardiovascular disease in adulthood. *Ann Epidemiol*. 2006;16:91–104.
30. Lawlor DA, Sterne JA, Tynelius P, Davey Smith G, Rasmussen F. Association of childhood socioeconomic position with cause-specific mortality in a prospective record linkage study of 1,839,384 individuals. *Am J Epidemiol*. 2006;164:907–915.

31. Lipowicz A, Koziel S, Hulanicka B, Kowalisko A. Socioeconomic status during childhood and health status in adulthood: the Wroclaw growth study. *J Biosoc Sci*. 2007;39:481–491.
32. Melchior M, Moffitt TE, Milne BJ, Poulton R, Caspi A. Why do children from socioeconomically disadvantaged families suffer from poor health when they reach adulthood? A life-course study. *Am J Epidemiol*. 2007;166:966–974.
33. Rudolph CW, Wells CL, Weller MD, Baltes BB. A meta-analysis of empirical studies of weight-based bias in the workplace. *J Vocat Behav*. 2009;74:1–10.
34. Shrewsbury V, Wardle J. Socioeconomic status and adiposity in childhood: a systematic review of cross-sectional studies 1990–2005. *Obesity*. 2008;16:275–284.
35. Parsons TJ, Power C, Logan S, Summerbell CD. Childhood predictors of adult obesity: a systematic review. *Int J Obes Relat Metab Disord.* 1999;23suppl 8:S1–S107.
36. Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol*. 2002;31:163–165.
37. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct EquModeling*. 2001;8:430–457.
38. VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology*. 2009;20:880–883.
39. Goetgeluk S, Vansteelandt S, Goetghebeur E. Estimation of controlled direct effects. *J R Stat Soc Series B Stat Methodol*. 2008;70:1049–1066.
40. van der Laan MJ, Petersen ML. Direct effect models. Int J *Biostatist.* 2008;4:Article 23.
41. Vansteelandt S. Estimating direct effects in cohort and case-control studies. *Epidemiology*. 2009;20:851–860.
42. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168:656–664.
43. Ranjit N, Diez-Roux AV, Shea S, Cushman M, Ni H, Seeman T. Socioeconomic position, race/ethnicity, and inflammation in the multi-ethnic study of atherosclerosis. *Circulation*. 2007;116:2383–2390.
44. Nazmi A, Victora CG. Socioeconomic and racial/ethnic differentials of C-reactive protein levels: a systematic review of population-based studies. *BMC Public Health.* 2007;7:212.
45. Thurston RC, Matthews KA. Racial and socioeconomic disparities in arterial stiffness and intima media thickness among adolescents. *Soc Sci Med*. 2009;68:807–813.
46. Cohen S, Janicki-Deverts D, Chen E, Matthews KA. Childhood socioeconomic status and adult health. *Ann N Y Acad Sci*. 2010;1186:37–55.
47. VanderWeele TJ. Direct and indirect effects for neighborhood-based clustered and longitudinal data. *Sociol MethodsRes*. 2010;38:515–544.

eAPPENDIX

Measurement of life-course SES

EFA with quartimin rotation was used to assess the dimensionality of our SES indicators (eTable 1). We used the root mean square error of approximation (RMSEA) and comparative fit index (CFI) to determine the smallest number of latent factors needed to explain the correlations among SES indicators. A RMSEA less than 0.06 and a CFI greater than 0.95 generally indicate a good fitting model.[1] A three factor model with a RMSEA=0.06 and CFI=0.97 was initially identified. The first factor was measured by five indicators of early-life SES, including father's and mother's educational attainment, father's occupational status, region of birth, and rural residence during childhood. Three indicators of early-life SES, including self-rated childhood SES, whether the respondent's family moved for financial reasons during childhood, and whether there was a period during childhood when the respondent's father was unemployed for several months measured the second factor. The third factor was measured by five adult SES indicators, including respondent's educational attainment, respondent's occupation, respondent's labor force status, household income, and household wealth. The three-indicator factor for early-life SES was deleted because the indicators were, based on face validity, poor measures of the latent factor. The final two-factor model had a reasonable fit based on the RMSEA (RMSEA=0.07) and a good fit based on the CFI (CFI=0.98).

We next used CFA with maximum likelihood estimation to output factor scores for early-life and adult SES. The results of the CFA, including standardized model coefficients and standard errors are shown in eFigure 1. Approximately 47% of participants were missing data on at least one indicator of early-life SES. Missing values were handled with full information maximum likelihood (FIML) using the expectation maximization algorithm. This method uses all available data to estimate model

parameters and has been shown, with ignorable missing data, to be unbiased and more efficient than other methods for handling missing data.[2] We conducted sensitivity analyses comparing results using FIML to listwise deletion; results were qualitatively similar although FIML estimates were more precise. Continuous early-life and adult SES scores were cut into quartiles for analysis. Analyses using tertiles of SES produced qualitatively similar results (not shown). Factor analyses were conducted in Mplus version 5.

Obtaining HRS data files and documentation

The Health and Retirement Study data files and documentation, including the longitudinal RAND HRS data file that was used to define measures for the majority of our analyses and the cross-wave tracker file that includes mortality information, are publicly available pursuant to registration through the University of Michigan's HRS website: http://hrsonline.isr.umich.edu/

SAS code for estimating total and controlled direct effects

In this section we provide SAS code for estimating the total and controlled direct effects of early-life SES on chronic disease that was not mediated by adult SES using the conventional regression approach to effect decomposition and the marginal structural model approach using inverse probability weights. In the models below: (1) the outcome variable, event, is a dichotomous variable indicating whether the participant reported the outcome, using coronary heart disease (CHD) as an example, during follow-up between 1992-2006; (2) early-life (C_4SES) and adult (A_4SES) SES scores were split into quartiles represented by categorical (e.g., C_4SES) or dummy variables (e.g., C_4SESc1-C_4SESc4); (3) measured confounders include dummy variables for sex (gen), birth cohort (cohort), race/ethnicity (race), and self-rated childhood health

(chealth); (4) risk factors include dummy variables for high blood pressure (hibp), body mass index (bmi), self-rated health (srh), diabetes (diab), functional limitations (hlthlw), current smoking (smoke), and alcohol consumption (alcohol); (5) HHIDPN is the participant identifier; and (6) lifeSES is the dataset with individual-level data on early-life SES, adult SES, confounders, risk factors, and CHD. The categorizations of confounders and risk-factors are defined in the text, as well as Table 1.

First, corresponding to Model 1 in the text, we estimated on the risk ratio scale the total effect of early-life SES on CHD by fitting a log-linear model regressing the event on quartiles of early-life SES (with the highest quartile as the reference category) and the set of confounders:

```
*Total effect of early-life SES on CHD;
proc genmod data=lifeSES descending;
class HHIDPN;
model event=C_4SESc2 C_4SESc3 C_4SESc4
genc2 cohortc2 cohortc3 cohortc4 racec2 chealthc2 chealthc3 chealthc4/dist=bin
link=log;
run;
```

Second, corresponding to Model 2 in the text, we estimated on the risk ratio scale the direct effect of early-life SES not mediated by adult SES by regressing the event on early-life SES, adult SES, and the set of confounders:

```
*Direct effect of early-life SES on CHD not through adult SES;
proc genmod data=lifeSES descending;
class HHIDPN;
model event=C_4SESc2 C_4SESc3 C_4SESc4
A_4SESc2 A_4SESc3 A_4SESc4
genc2 cohortc2 cohortc3 cohortc4 racec2 chealthc2 chealthc3 chealthc4/dist=bin
link=log;
run;
```

Third, corresponding to Model 3 in the text, we estimated on the risk ratio scale the direct effect of early-life SES not mediated by adult SES after additionally including adult risk factors as covariates. Note in the model below that we changed the distribution from binomial to poisson because the binomial model failed to converge:[3]

```
*Direct effect of early-life SES on CHD not through adult SES after accounting for adult
risk factors;
proc genmod data=lifeSES descending;
class HHIDPN;
model event=C_4SESc2 C_4SESc3 C_4SESc4
A_4SESc2 A_4SESc3 A_4SESc4
genc2 cohortc2 cohortc3 cohortc4 racec2 chealthc2 chealthc3 chealthc4
hibpc2 bmic2 bmic3 srhc2 srhc3 srhc4 srhc5 diabc2 hlthlwc2 smokec2 alcoholc2
alcoholc3 alcoholc4/dist=poisson link=log;
run;
```

Fourth, we estimated the controlled direct effect of early-life SES on CHD by a marginal structural model, as described by Model 4 in the text. Recall from the text that potential confounding by the sets of confounders and risk factors was accounted for by fitting the marginal structural model with stabilized inverse probability weights of the form $W = w_i^A * w_i^M$, where the weight w_i^A accounted for measured confounding of the relation between early-life SES and CHD and the weight w_i^M accounted for measured confounding of the relation between adult SES and CHD.

In order to calculate the weight w_i^A, corresponding to Model 5 in the text, we used two ordinal logistic regression models. The first model, which is used to calculate the numerator of w_i^A, is an empty model with quartiles of early-life SES as the outcome; including the predicted probabilities of early-life SES in the numerator of w_i^A stabilizes the weight. The second model, which is used to calculate the denominator of w_i^A, regresses categories of early-life SES on the set of confounders:

*Model for calculating the numerator of weight w_i^A;
```
proc genmod data=lifeSES;
class HHIDPN;
model C_4SES=/d=multinomial link=clogit;
output out=pred_n1 p=n1;
run;
```

*Model for calculating the denominator of weight w_i^A;
```
proc genmod data=lifeSES;
class HHIDPN;
model C_4SES=genc2 cohortc2 cohortc3 cohortc4 racec2 chealthc2 chealthc3
chealthc4/d=multinomial link=clogit;
output out=pred_d1 p=d1;
run;
```

From the numerator and denominator models for w_i^A above, we output two new datasets, pred_n1 and pred_d1, respectively. These datasets include for each individual the cumulative predicted probabilities of early-life SES, n1 and d1, arranged by the ordinal level value _level_. In order to calculate the quartile-specific probabilities for the numerator and denominator, we sorted the cumulative predicted probabilities of early-life SES, transposed them into wide datasets pred_n2 and pred_d2, respectively, and subtracted to calculate the predicted probability of each quartile of early-life SES.

*Sort dataset with cumulative predicted probabilities of early-life SES, represented by n1, by HHIDPN and _level_, from the numerator model for w_i^A (pred_n1);
```
proc sort data=pred_n1;
by HHIDPN _level_;
run;
```

*Transpose to create wide dataset of cumulative probabilities, pred_n2, from pred_n1;
```
proc transpose DATA=pred_n1 OUT=pred_n2 (drop=_name_ _label_);
by HHIDPN;
var n1;
id _level_;
run;
```

*Calculate the predicted probabilities of each quartile of early-life SES, n11-n14, for the numerator of w_i^A;

```
data pred_n2;
set pred_n2;
n11=_1;
n12=(_2-_1);
n13=(_3-_2);
n14=(1.0-_3);
run;
```

*****Sort dataset with cumulative predicted probabilities of early-life SES, represented by d1, by HHIDPN and _level_, from the numerator model for w_i^A (pred_d1);

```
proc sort data=pred_d1;
by HHIDPN _level_;
run;
```

```
*Transpose to create wide dataset of cumulative probabilities, pred_d2, from pred_d1;
proc transpose DATA=pred_d1 OUT=pred_d2 (drop=_name_ _label_);
by HHIDPN;
var d1;
id _level_;
run;
```

*Calculate the predicted probabilities of each quartile of early-life SES, d11-d14, for the denominator of w_i^A;

```
data pred_d2;
set pred_d2;
d11=_1;
d12=(_2-_1);
d13=(_3-_2);
d14=(1.0-_3);
run;
```

We then sorted and merged datasets with the predicted probabilities of early-life SES for the numerator and denominator of w_i^A with the lifeSES dataset containing covariate information, calculated the predicted probability of early-life SES based on the quartile of early-life SES actually reported, and divided the numerator and denominator of w_i^A, num1 and den1, respectively, to obtain the stabilized weight, w_i^A:

*Sort datasets with the predicted probabilities of early-life SES for the numerator and denominator of w_i^A with the lifeSES dataset containing covariate information;

```
proc sort data=lifeSES_nomiss; by HHIDPN;
```

```
proc sort data=pred_n2; by HHIDPN;
proc sort data=pred_d2; by HHIDPN; run;
```

*Merge datasets and calculate the predicted probability of early-life SES based on the quartile of early-life SES actually reported and divide the numerator and denominator of w_i^A, num1 and den1, respectively, to obtain the stabilized weight, w_i^A:

```
data lifeSES_nomiss;
set lifeSES_nomiss;
merge lifeSES_nomiss pred_n2 pred_d2;
by HHIDPN;
if C_4SES=4 then do; num1=n14; den1=d14; end;
if C_4SES=3 then do; num1=n13; den1=d13; end;
if C_4SES=2 then do; num1=n12; den1=d12; end;
if C_4SES=1 then do; num1=n11; den1=d11; end;
W1=round(num1/den1,.001);
run;
```

In the following steps, we calculated a weight that accounts for measured confounding of the relation between adult SES and CHD, w_i^M, accounting for the covariates in Model 6 of the text, using an analogous series of commands:

*Model for calculating the numerator of weight w_i^M;

```
proc genmod data=lifeSES_nomiss;
class HHIDPN;
model A_4SES=C_4SESc2 C_4SESc3 C_4SESc4/d=multinomial link=clogit;
output out=pred_n3 p=n3;
run;
```

*Model for calculating the denominator of weight w_i^M;

```
proc genmod data=lifeSES_nomiss;
class HHIDPN;
model A_4SES=C_4SESc2 C_4SESc3 C_4SESc4
genc2 cohortc2 cohortc3 cohortc4 racec2 chealthc2 chealthc3 chealthc4
hibpc2 bmic2 bmic3 srhc2 srhc3 srhc4 srhc5
diabc2 hlthlwc2 smokec2 alcoholc2 alcoholc3 alcoholc4/d=multinomial link=clogit;
output out=pred_d3 p=d3;
run;
```

*Sort dataset with cumulative predicted probabilities of adult SES, represented by n3, by HHIDPN and _level_, from the numerator model for w_i^M (pred_n3);

```
proc sort data=pred_n3;
by HHIDPN _level_;
```

```
run;

*Transpose to create wide dataset of cumulative probabilities, pred_n4, from pred_n3;
proc transpose DATA=pred_n3 OUT=pred_n4 (drop=_name_ _label_);
by HHIDPN;
var n3;
id _level_;
run;

*Calculate the predicted probabilities of each quartile of adult SES, n21-n24, for the
numerator of w_i^M ;
data pred_n4;
set pred_n4;
n21=_1;
n22=(_2-_1);
n23=(_3-_2);
n24=(1.0-_3);
run;

*Sort dataset with cumulative predicted probabilities of adult SES, represented by d3, by
HHIDPN and _level_, from the denominator model for w_i^M (pred_d3);
proc sort data=pred_d3;
by HHIDPN _level_;
run;

*Transpose to create wide dataset of cumulative probabilities, pred_d4, from pred_d3;
proc transpose DATA=pred_d3 OUT=pred_d4 (drop=_name_ _label_);
by HHIDPN;
var d3;
id _level_;
run;

*Calculate the predicted probabilities of each quartile of adult SES, d21-d24, for the
denominator of w_i^M ;
data pred_d4;
set pred_d4;
d21=_1;
d22=(_2-_1);
d23=(_3-_2);
d24=(1.0-_3);
run;

*Sort datasets with the predicted probabilities of early-life SES for the numerator and
denominator of w_i^A with the lifeSES dataset containing covariate information;
proc sort data=lifeSES_nomiss; by HHIDPN; run;
proc sort data=pred_n4; by HHIDPN; run;
proc sort data=pred_d4; by HHIDPN; run;

*Merge datasets and calculate the predicted probability of early-life SES based on the
```

quartile of adult SES actually reported and divide the numerator and denominator of w_i^M , num2 and den2, respectively, to obtain the stabilized weight, w_i^M ;

```
data lifeSES_nomiss;
set lifeSES_nomiss;
merge lifeSES_nomiss pred_n4 pred_d4;
by HHIDPN;
if A_4SES=4 then do; num2=n24; den2=d24; end;
if A_4SES=3 then do; num2=n23; den2=d23; end;
if A_4SES=2 then do; num2=n22; den2=d22; end;
if A_4SES=1 then do; num2=n21; den2=d21; end;
W2=round(num2/den2,.001);
run;
```

Next, we multiply w_i^A and w_i^M to obtain the stabilized inverse probability weight, W:

```
data lifeSES_nomiss;
set lifeSES_nomiss;
W=W1*W2;
run;
```

Finally, we estimated the controlled direct effect of early-life SES on CHD by a stabilized

inverse probability weighted marginal structural, corresponding to Model 4 in the text:

```
*Controlled direct effect of early-life SES on CHD by stabilized inverse probability
weighted marginal structural model;
proc genmod data=lifeSES descending;
class HHIDPN;
model event=C_4SESc2 C_4SESc3 C_
4SESc4 A_4SESc2 A_4SESc3 A_4SESc4/dist=bin link=log;
weight W;
run;
```

REFERENCES (eAPPENDIX)

1. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling.* 1999;6:1-55.

2. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling.* 2001;8:430-457.

3. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am. J. Epidemiol.* 2005;162:199-200.

eTable 1. Summary of indicators used to measure SES in early-life and adulthood, Health and Retirement Study

Indicator[a]	Year[b]	Description	Latent factor[c]
		Indicators of early-life SES	
Father's educational attainment (years)	1992	4 indicators (0-7, 8-11, 12, >12)	Early-life SES
Mother's educational attainment (years)	1992	4 indicators (0-7, 8-11, 12, >12)	Early-life SES
Father's occupation	1998	Dichotomous (manual/unskilled service, professional/white collar)	Early-life SES
Birth in the southern US	1992	Dichotomous (yes/no)	Early-life SES
Rural residence during childhood	1992	Dichotomous (yes/no)	Early-life SES
Self rated childhood SES	1998	3 indicators (varied/poor, about average, pretty well off financially)	Not included in final model
Family moved for financial reasons	1998	Dichotomous (yes/no)	Not included in final model
Father unemployed for several months	1998	3 indicators (no, yes/never lived with father)	Not included in final model
		Indicators of adult SES	
Respondent's educational attainment	1992	4 indicators (<HS, HS/GED, some college, college+)	Adult SES
Respondent's longest job occupation	1992	Dichotomous (manual/unskilled service, professional/white collar)	Adult SES
Respondent's labor force status	1992	Dichotomous (works FT/PT/retired, unemployed/disabled/not in labor force)	Adult SES
Household income	1992	Split into quartiles (indicators)	Adult SES
Household wealth	1992	Split into quartiles (indicators)	Adult SES
Adult height	1992	Split into quartiles (indicators)	Not included in final model

[a] Hypothesized indicators of SES selected from the Health and Retirement Study
[b] Year of assessment
[c] Indicates the latent factor measured by each indicator in the final two-factor model

eTable 2. Coefficients, standard errors (SE), and p-values from model (Model 2) regressing outcomes on main effects of early-life and adult SES and cross-product terms representing interaction between early-life and adult SES, accounting for adult SES and confounders

Life-course socioeconomic status	CHD: Regression adjustment for adult SES and confounders (Model 2)[b]			Diabetes: Regression adjustment for adult SES and confounders (Model 2)[b]			Stroke: Regression adjustment for adult SES and confounders (Model 2)[b]		
	Beta	**SE**	**p-value**	**Beta**	**SE**	**p-value**	**Beta**	**SE**	**p-value**
Early-life SES[a]									
A2	0.130	0.088	0.138	-0.046	0.131	0.726	-0.184	0.215	0.394
A3	0.055	0.165	0.736	0.495	0.178	0.006	0.213	0.324	0.511
A4 (lowest SES)	-0.245	0.291	0.400	0.321	0.279	0.251	-0.142	0.573	0.805
Adult SES[a]									
M2	0.157	0.088	0.075	0.213	0.119	0.074	0.354	0.181	0.051
M3	0.232	0.151	0.123	0.185	0.208	0.374	0.599	0.262	0.022
M4 (lowest SES)	0.512	0.367	0.164	0.371	0.522	0.477	0.349	0.969	0.719
Early-life * Adult SES									
A2*M2	-0.226	0.131	0.084	0.090	0.180	0.618	-0.680	0.318	0.033
A2*M3	-0.334	0.185	0.070	0.126	0.254	0.620	0.100	0.346	0.772
A2*M4	-0.540	0.407	0.185	0.065	0.561	0.907	0.305	1.020	0.765
A3*M2	-0.014	0.193	0.941	-0.789	0.234	0.001	-0.216	0.382	0.571
A3*M3	-0.159	0.225	0.480	-0.287	0.274	0.294	-0.347	0.415	0.403
A3*M4	-0.275	0.405	0.497	-0.260	0.552	0.638	0.170	1.022	0.868
A4*M2	0.129	0.332	0.699	-0.212	0.338	0.531	-0.319	0.674	0.636
A4*M3	0.054	0.335	0.873	-0.141	0.355	0.692	-0.127	0.639	0.843
A4*M4	0.055	0.467	0.906	-0.006	0.589	0.991	0.617	1.122	0.582

[a] Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles of early-life and adult SES

[b] Adjusted for early-life SES, adult SES, and set of confounders (i.e., birth cohort, sex, race/ethnicity, and self-rated childhood health)

eTable 3. Coefficients, standard errors (SE), and p-values from model (Model 3) regressing outcomes on main effects of early-life and adult SES and cross-product terms representing interaction between early-life and adult SES, accounting for adult SES, confounders, and risk factors

Life-course socioeconomic status	CHD: Regression adjustment for adult SES, confounders and risk factors (Model 3)[b]			Diabetes: Regression adjustment for adult SES, confounders and risk factors (Model 3)[b]			Stroke: Regression adjustment for adult SES, confounders and risk factors (Model 3)[b]		
	Beta	**SE**	**p-value**	**Beta**	**SE**	**p-value**	**Beta**	**SE**	**p-value**
Early-life SES[a]									
A2	0.089	0.102	0.382	-0.072	0.140	0.609	-0.183	0.221	0.408
A3	-0.055	0.189	0.771	0.319	0.201	0.112	0.072	0.337	0.831
A4 (lowest SES)	-0.310	0.322	0.336	0.084	0.312	0.788	-0.123	0.590	0.835
Adult SES[a]									
M2	-0.047	0.103	0.650	0.065	0.130	0.617	0.132	0.190	0.488
M3	-0.106	0.177	0.550	-0.054	0.228	0.812	0.182	0.279	0.515
M4 (lowest SES)	0.278	0.453	0.539	0.141	0.584	0.809	0.086	1.010	0.932
Early-life * Adult SES									
A2*M2	-0.101	0.151	0.505	0.080	0.195	0.683	-0.602	0.327	0.066
A2*M3	-0.186	0.214	0.384	0.161	0.276	0.559	0.202	0.362	0.576
A2*M4	-0.658	0.496	0.185	-0.003	0.626	0.996	0.049	1.063	0.963
A3*M2	0.146	0.224	0.515	-0.681	0.259	0.009	-0.049	0.398	0.902
A3*M3	0.046	0.261	0.860	-0.151	0.303	0.618	-0.116	0.435	0.790
A3*M4	-0.316	0.494	0.523	-0.201	0.619	0.746	0.038	1.064	0.972
A4*M2	0.316	0.371	0.395	0.058	0.376	0.877	-0.231	0.693	0.739
A4*M3	0.242	0.376	0.519	0.072	0.394	0.856	-0.010	0.661	0.988
A4*M4	-0.057	0.554	0.919	0.086	0.659	0.897	0.309	1.165	0.791

[a] Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles of early-life and adult SES

[b] Adjusted for early-life SES, adult SES, set of confounders (i.e., birth cohort, sex, race/ethnicity, and self-rated childhood health), and set of risk factors (i.e., high blood pressure, BMI, self-rated health, diabetes, functional limitations, current smoking, alcohol consumption)

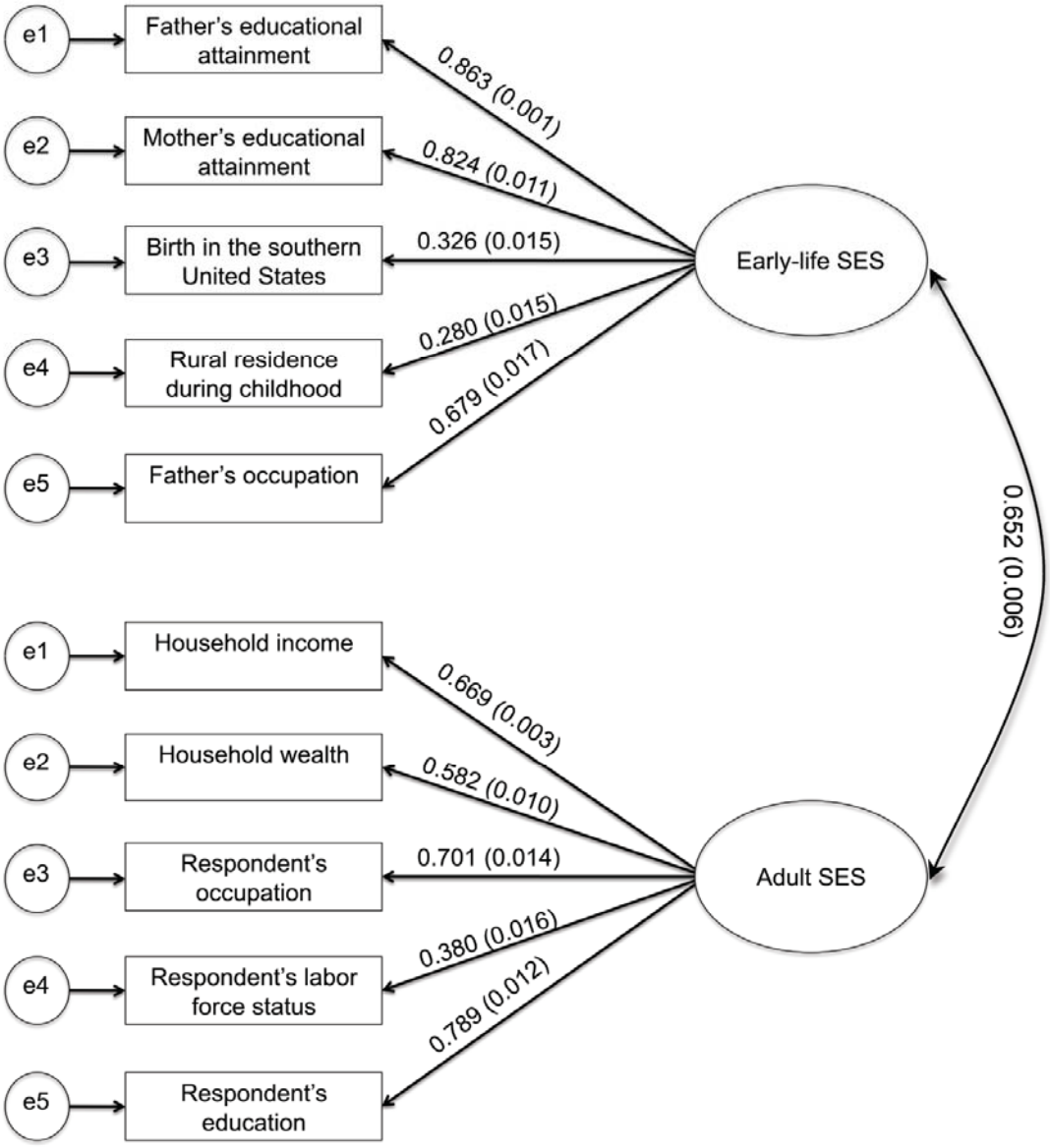
eTable 4. Coefficients, standard errors (SE), and p-values from marginal structural model (Model 4) regressing outcomes on main effects of early-life and adult SES and cross-product terms representing interaction between early-life and adult SES, accounting for adult SES, confounders, and risk factors by stabilized inverse probability weights

Life-course socioeconomic status	**CHD: Marginal structural model** (Model 4)[b]			**Diabetes: Marginal structural model** (Model 4)[b]			**Stroke: Marginal structural model** (Model 4)[b]		
	Beta	**SE**	**p-value**	**Beta**	**SE**	**p-value**	**Beta**	**SE**	**p-value**
Early-life SES[a]									
A2	0.156	0.084	0.065	-0.088	0.123	0.475	-0.275	0.218	0.208
A3	0.198	0.146	0.176	0.515	0.167	0.002	0.755	0.247	0.002
A4 (lowest SES)	0.153	0.229	0.503	0.258	0.289	0.373	-0.148	0.618	0.811
Adult SES[a]									
M2	0.020	0.092	0.827	0.008	0.120	0.949	0.171	0.188	0.362
M3	-0.010	0.167	0.952	-0.022	0.217	0.920	0.220	0.311	0.479
M4 (lowest SES)	0.677	0.318	0.034	0.362	0.518	0.484	0.797	0.746	0.285
Early-life * Adult SES									
A2*M2	-0.161	0.131	0.218	0.286	0.175	0.102	-0.337	0.318	0.290
A2*M3	-0.284	0.200	0.157	0.096	0.263	0.715	0.339	0.393	0.387
A2*M4	-0.928	0.376	0.014	0.075	0.559	0.894	-0.820	0.881	0.352
A3*M2	0.067	0.178	0.707	-0.630	0.224	0.005	-0.546	0.323	0.091
A3*M3	-0.113	0.224	0.614	-0.246	0.275	0.372	-0.555	0.396	0.161
A3*M4	-0.879	0.357	0.014	-0.446	0.547	0.415	-0.967	0.789	0.220
A4*M2	-0.081	0.283	0.776	-0.004	0.349	0.990	-0.234	0.737	0.751
A4*M3	0.001	0.290	0.998	-0.109	0.371	0.770	0.460	0.698	0.509
A4*M4	-0.663	0.392	0.090	-0.068	0.591	0.909	0.093	0.966	0.923

[a] Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles of early-life and adult SES

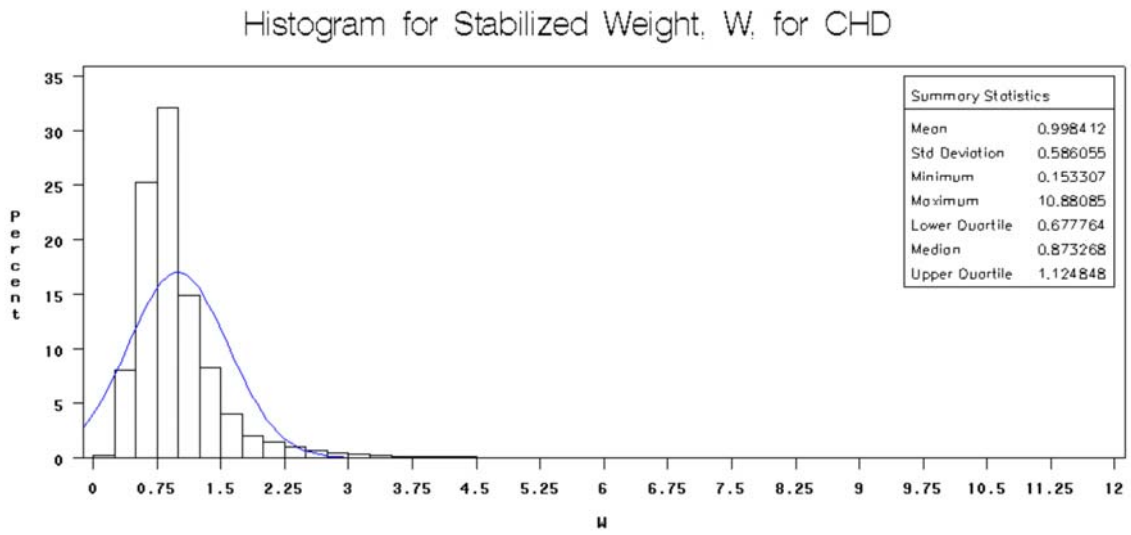
[b] Stabilized inverse probability weights used to account for potential confounding by the sets of confounders and risk factors

eFigure 1. Standardized model coefficients and standard errors (in parentheses) from confirmatory factor analysis for measurement of early-life and adult SES; n=9,760 Health and Retirement Study enrollees followed from 1992-2006[a]

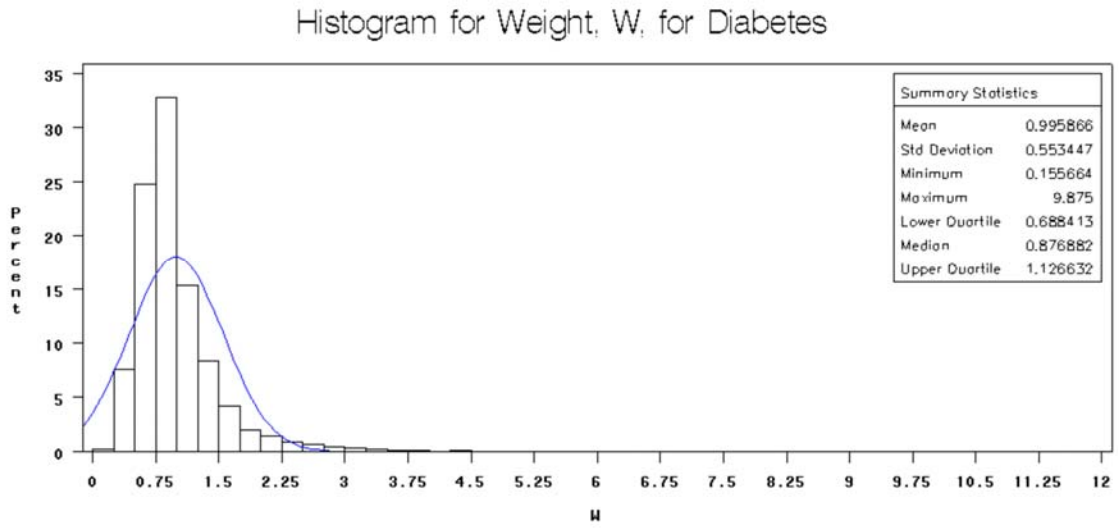


[a] Coefficients and standard errors standardized using both latent variable and observed variable variances

eFigure 2. Histogram and summary statistics showing distribution of stabilized weight, W, for heart disease, n=8308



eFigure 3. Histogram and summary statistics showing distribution of stabilized weight, W, for diabetes, n=8267



eFigure 4. Histogram and summary statistics showing distribution of stabilized weight, W, for stroke, n=9041