Instrumental-Variables Simultaneous Equations Model of Physical Activity and Body Mass Index

The Coronary Artery Risk Development in Young Adults (CARDIA) Study

Katie A. Meyer, David K. Guilkey, Hsiao-Chuen Tien, Catarina I. Kiefe, Barry M. Popkin, and Penny Gordon-Larsen*

* Correspondence to Dr. Penny Gordon-Larsen, Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, 137 East Franklin Street, 6th Floor, Campus Box #8120, Chapel Hill, NC 27514 (e-mail: pglarsen@unc.edu).

Initially submitted June 12, 2015; accepted for publication January 11, 2016.

We used full-system-estimation instrumental-variables simultaneous equations modeling (IV-SEM) to examine physical activity relative to body mass index (BMI; weight (kg)/height (m)^2) using 25 years of data (1985/1986 to 2010/2011) from the Coronary Artery Risk Development in Young Adults (CARDIA) Study ($n = 5{,}115$; ages 18–30 years at enrollment). Neighborhood environment and sociodemographic instruments were used to characterize physical activity, fast-food consumption, smoking, alcohol consumption, marriage, and childbearing (women) and to predict BMI using semiparametric full-information maximum likelihood estimation to control for unobserved time-invariant and time-varying residual confounding and differential measurement error through model-derived discrete random effects. Comparing robust-variance ordinary least squares, random-effects regression, fixed-effects regression, single-equation-estimation IV-SEM, and full-system-estimation IV-SEM, estimates from random- and fixed-effects models and the full-system-estimation IV-SEM were unexpectedly similar, despite the lack of control for residual confounding with the random-effects estimator. Ordinary least squares tended to overstate the significance of health behaviors in BMI, while results from single-equation-estimation IV-SEM were notably different, revealing the impact of weak instruments in standard instrumental-variable methods. Our robust findings for fixed effects (which does not require instruments but has a high cost in lost degrees of freedom) and full-system-estimation IV-SEM (vs. standard IV-SEM) demonstrate potential for a full-system-estimation IV-SEM method even with weak instruments.

body mass index; endogeneity; epidemiologic methods; fixed effects; health behaviors; instrumental variables; semiparametric methods; simultaneous equations

Abbreviations: BMI, body mass index; CARDIA, Coronary Artery Risk Development in Young Adults; CI, confidence interval; FIML, full-information maximum likelihood; GMM, generalized method of moments; IV, instrumental variable; IV-SEM, instrumental-variables simultaneous equations model/modeling; OLS, ordinary least squares.

A critical challenge in observational studies of the influence of health behaviors, such as physical activity, on body weight (1, 2) is the potential for residual confounding due to omitted variables or differential measurement error (referred to in econometrics as endogeneity) (3–8). The assessment of diet and physical activity is known to be susceptible to error, which may be differential by body weight (9, 10); more generally, confounding may reflect difficult-to-measure innate characteristics, such as underlying health consciousness (3, 11).

Standard epidemiologic analyses of health behaviors and body mass index (BMI) are susceptible to residual confounding due to observed variables and differential measurement error. In the current study, our goal was to estimate the effect of physical activity on BMI, while accounting for these biases. We addressed the potential for residual confounding and differential measurement error in our analysis with an instrumental-variables simultaneous equations modeling (IV-SEM) approach commonly

used in econometric studies using longitudinal data (8, 11–13).

Econometric approaches to causal inference have been employed in the epidemiologic literature (14–18) but have typically been limited to single-equation systems (11), such as the role of a specific genetic variant in an outcome (19), or systems models fitted using single-equation methods (11), such as 2-stage least squares (20). Our approach has several enhancements over standard instrumental-variable (IV) methods, including joint estimation of an entire system of equations that accounts for unmeasured confounding and differential measurement error in the analysis of multiple BMI risk factors, and the use of semiparametric and nonlinear estimation. Our method has been shown to perform better than linear IV methods even in the presence of weak instruments—a well-known challenge for IV methods (21, 22). For these reasons, we consider our approach a valuable addition to the IV methodological framework considered by epidemiologists.

We estimated the effect of physical activity on BMI using 25 years of data from an established prospective cohort study, the Coronary Artery Risk Development in Young Adults (CARDIA) Study, with clinic-assessed, time-varying measures of weight and height, weight-related health behaviors, and other relevant variables, as well as an extensive set of community-level variables hypothesized to influence BMI-related risk factors but not BMI directly, which served as IVs (11, 23). These data provided the necessary components for an IV-SEM with which to estimate the effect of physical activity on BMI while accounting for other health behaviors, including diet, smoking, alcohol consumption, and marital status, as well as residual confounding and differential measurement error. As a secondary aim, we examined the extent to which our estimates differed from those obtained via other modeling approaches. We therefore compared estimates from our model with those from standard ordinary least squares (OLS) with robust variance, longitudinal random-effects and fixed-effects regression models, and single-equation-estimation IV-SEM.

METHODS

CARDIA sample

CARDIA is a multicenter, longitudinal study of cardiometabolic risk factors (24). The study began in 1985–1986 with 5,115 black and white adults aged 18–30 years sampled from 4 US metropolitan areas (Birmingham, Alabama; Chicago, Illinois; Minneapolis, Minnesota; and Oakland, California). Participant home addresses were geocoded at baseline (year 0) and at years 7, 10, 15, 20, and 25 of follow-up (respective retention in survivors: 81%, 79%, 74%, 72%, and 72%). The CARDIA protocol was approved by the institutional review board at each field center, and every participant provided informed written consent.

Individual-level measures

In standardized surveys, participants provided extensive demographic and socioeconomic information, including age, sex, race, education, income, marital status, the ages of any children, and, for women, pregnancy status. Participants reported their engagement in 13 physical activities, including walking, running, and cycling, from which activity-specific and total activity intensity scores were created (25). Fast-food consumption, smoking status, and alcohol consumption were assessed in all years; an interviewer-administered diet history was included at years 0 (baseline), 7, and 20 (26). At each examination, height and weight were measured by trained study staff to the nearest 0.5 cm and 0.2 kg, respectively. BMI was calculated as weight (kg)/height $(m)^2$.

Neighborhood-level measures

Neighborhood measures temporally and geographically linked to participant home addresses served as instruments for identification of estimated effects of health behaviors on BMI. Briefly, neighborhood measures included the presence of food stores and restaurants, physical activity facilities, and parks; consumer price data; and features of the road network (details are provided in the Web Appendix, available at http://aje.oxfordjournals.org/). US Census data included population-level educational attainment and income for the census tract corresponding to participants' home addresses at the time of the examination.

Statistical analysis

Figure 1 is an abbreviated causal diagram of our statistical model. We sought to estimate the effects of physical activity and other health behaviors (the vector ***HB***) on BMI. We considered the potential for time-invariant and time-varying residual confounding and differential measurement error, μ, to cause bias in effect estimates, and we used a full-system-estimation IV-SEM to estimate the effects of health behaviors on BMI by using a set of IVs, *V*, to first identify variation in health behaviors. A complete list of model variables is shown in the Web Appendix and Web Table 1.

Econometricians refer to the differential error that concerns us, μ, such as residual confounding, as "endogeneity due to unobserved heterogeneity," where "unobserved heterogeneity" reflects individual heterogeneity in the outcome (here, BMI) that 1) is not explained by independent variables (predictors already included in the regression model) and 2) is correlated with independent variables. "Endogeneity" refers to variables that are determined within the model (i.e., are recipients of an inward-pointing arrow on the causal diagram); in contrast, "exogenous" variables are not determined by other variables in the system (i.e., have only outward-pointing arrows). Variables *V* and *X* in Figure 1 are exogenous. Both endogenous and exogenous variables influence other system variables. Here, endogenous health behaviors influence BMI; exogenous IVs, *V*, influence health behaviors but not BMI directly. There may be many instances where formal endogeneity will not fit within our causal framework, such as reverse causation, whereby BMI influences health behaviors. We acknowledge that the exact sources of unobserved heterogeneity quantified by our model-based approach cannot be distinguished.

Previous publications provide more details of our approach (27–29). Additional details on system equations can be found

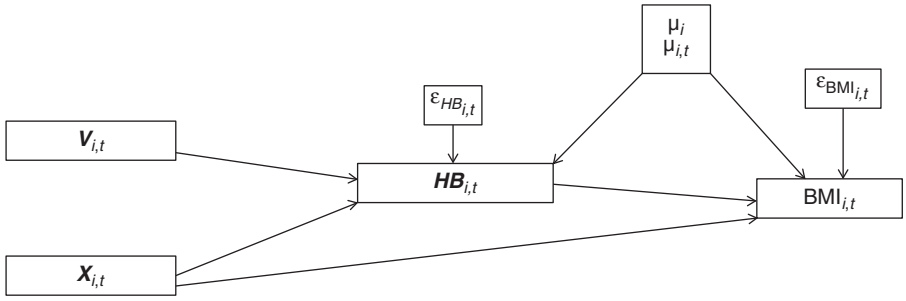

Figure 1. Causal diagram for a full-system-estimation instrumental-variables simultaneous equations model created to examine physical activity in relation to body mass index (BMI). Time-varying BMI ($BMI_{i,t}$) was predicted from time-varying physical activity and other health behaviors ($\boldsymbol{HB}_{i,t}$). $\boldsymbol{V}_{i,t}$ represents a vector of instrumental variables that are associated with BMI only through their effect on health behavior. $\boldsymbol{X}_{i,t}$ represents a vector of noninstrumental exogenous variables that may influence health behaviors and BMI. First-stage equations estimate HB from *V* and *X*, with distinct equations for each health behavior. After identification of HB from exogenous variables (*V* and *X*), BMI is predicted from the set of health behaviors and noninstrumental exogenous covariates *X*. Random error associated with each health behavior and BMI is shown by $\varepsilon_{i,t}$. Endogeneity due to unobserved heterogeneity (residual confounding or differential measurement error) is reflected by μ, with subscripts for time-invariant and time-varying components.

in the Web Appendix. Briefly, in a series of first-stage regression equations, we used a set of IVs (and other exogenous variables, *X* (in Figure 1), distinguished from IVs by their possible direct influence on BMI) to predict physical activity and other health behaviors, after which we estimated the effect of health behaviors on BMI, accounting for endogeneity due to unobserved heterogeneity (reflecting residual confounding and differential measurement error).

In addition to the estimated effect of total physical activity on BMI, we estimated the effects of specific types of physical activities most relevant for our neighborhood environment variables: 1) walking (controlling for all 12 nonwalking physical activities) and 2) walking, running, and cycling (controlling for all 10 other physical activities). Physical activity component variables were obtained by summing intensity scores over the relevant activities. We hypothesized that walking, along with running and cycling, would be most affected by neighborhood features—such as the road network—as opposed to other activities, such as swimming, which may require dedicated facilities.

We modeled other BMI risk factors, including fast-food consumption, smoking, and alcohol consumption, as well as marital status and, among women, childbearing, that we considered endogenous variables both associated with physical activity and predictive of BMI. Our full-system-estimation IV-SEM allowed us to account for confounding due to unobserved selectivity of individuals engaging in these activities, as well as residual confounding and differential measurement error in endogenous covariates themselves. Models were fitted separately for men and women to allow us to control for the selectivity of women who were pregnant at the time of each examination and control for childbearing history.

We considered, but did not explicitly control for, overall diet quality in our final models because diet quality (modeled as a score reflecting comprehensive food consumption) was not predictive of BMI in multivariable-adjusting models. Further, diet was assessed in only 3 examination periods, and this would have limited the analysis to a fraction of the study data. Any unobserved heterogeneity due to diet will be controlled in IV-SEM models and, to the extent that diet is stable within-person, in fixed-effects regression.

The general specification for the set of first-stage equations is shown below.

$$\mathrm{HB}_{it} = \gamma V_{it} + \delta X_{it} + \mu_i + \mu_{it} + \varepsilon_{it} \qquad (1)$$

Health behavior (HB) was modeled as a function of strictly exogenous explanatory variables, not influenced by other components within our model system, including a vector of IVs (***V***)—which influence BMI through health behaviors, but do not directly influence BMI—and a vector of other exogenous variables (***X***)—which may influence health behaviors as well as BMI. Health behaviors, exogenous variables, and random error (ε) were time-varying. Estimated effects of exogenous variables, *V* and *X*, are reflected by the regression coefficients γ and δ, respectively, for the difference in health behavior per unit change in *V* or *X*. In addition, our model included time-invariant and time-varying error components reflective of residual confounding and differential measurement error (μ). Each health behavior (endogenous BMI predictor) was modeled in a set of first-stage equations, with the full set repeated for each of the 3 physical activity specifications among men and women. Each first-stage equation was "overidentified," meaning the number of instruments exceeded the number of health behaviors, and comprised the same set of independent variables (11).

The final equation in the model estimated BMI from physical activity and other health behaviors (HB) and a set of exogenous covariates (*X*) that were considered to be associated with BMI (e.g., age, race, education, and income) beyond that accounted for by our set of modeled health behaviors. Note that this is the same set of *X* exogenous variables included in the first-stage equations.

$$\mathrm{BMI}_{it} = \beta \mathrm{HB}_{it} + \gamma X_{it} + \mu_i + \mu_{it} + \varepsilon_{it} \qquad (2)$$

Table 1. Key Attributes That Differentiate Regression Approaches

Attribute	OLS[a]	Random Effects[b]	Models That Control for Endogeneity Due to Unobserved Heterogeneity[c]			
			Fixed Effects[d]	IV Models[e]		
				Single-Equation System IV Regression[f]	GMM IV-SEM[g]	FIML IV-SEM[h]
Controls for time-invariant random error across individuals (individual-level unobserved heterogeneity) and uses error structure in parameter estimation (improved efficiency)	No	Yes	No	No	No	Yes
Controls for time-invariant endogeneity due to unobserved heterogeneity	No	No	Yes	Yes	Yes	Yes
Controls for time-varying endogeneity due to unobserved heterogeneity	No	No	No	Yes	Yes	Yes
Multiple-equation system, single-equation IV estimator	No	No	No	No	Yes	No
Multiple-equation system, full-system IV estimator: uses correlation among endogenous independent variables in parameter estimation (increased precision as compared with standard IV methods)	No	No	No	No	No	Yes
Nonlinear IV estimator (including discrete factor model for unobserved heterogeneity): less sensitive to weak IVs, as compared with other IV methods	No	No	No	No	No	Yes
Low power due to loss of degrees of freedom, compared with other approaches that control for endogeneity due to unobserved heterogeneity	No	No	Yes	No	No	No
Loss of precision if little within-person exposure variability	No	No	Yes	No	No	No
Allows estimation of the effects of time-invariant predictor variables, as compared with fixed effects	No	No	No	Yes	Yes	Yes
Requires valid IVs	No	No	No	Yes	Yes	Yes
Allows estimation of the effects of multiple endogenous variables	No	No	Yes	No	Yes	Yes
Use of multiple instruments allows testing of identification and can improve precision	No	No	No	No	Yes	Yes

Abbreviations: FIML, full-information maximum likelihood; GMM, generalized method of moments; IV, instrumental variable; IV-SEM, instrumental-variables simultaneous equations model; OLS, ordinary least squares.

[a] OLS regression models using the Stata -`regress`- command (StataCorp LP, College Station, Texas) with the robust variance option.

[b] Longitudinal random-effects regression models using the Stata -`xtreg`- command with the "`re`" option.

[c] Models can be distinguished broadly as those that directly address endogeneity from unobserved heterogeneity stemming from, for example, residual confounding and differential measurement error, including fixed-effects regression and the IV models, as opposed to those that assume no differential unobserved heterogeneity conditional on covariates included in the regression model, such as OLS and random-effects regression.

[d] Fixed-effects regression estimates for within-person differences from longitudinal data, obtained using the Stata -`xtreg`- command with the "`fe`" option.

[e] Among models that address endogeneity from unobserved heterogeneity, we compared 3 IV approaches. We included single-equation system IV regression for comparative purposes, given its prevalence in the epidemiologic literature, though we do not present results from this approach.

[f] Attributes hold for all IV regression, including single-equation systems, such as a model of a single IV for a single (endogenous) predictor variable. An example is instrumenting smoking behavior with a single genetic variant to predict lung cancer.

[g] IV methods for a system of equations that are estimated as a series of single equations. These simultaneous equations modeling approaches, including 2-stage least squares or GMM estimators, are available in standard statistical software, such as Stata's -`ivregress`- with the "`gmm`" or "`2sls`" option.

[h] IV methods for a system of equations that are estimated jointly. These approaches include FIML and 3-stage least squares, and are distinguished from single-equation-estimation approaches by their accounting for error correlation among equations.

Estimated effects of endogenous health behaviors (HB) and exogenous (X) independent variables are reflected by the regression coefficients β and γ, respectively, for the difference in BMI per unit change in the independent variable. The complete set of system equations is shown in the Web Appendix.

As with all IV approaches, our method relies on having a set of IVs that are 1) substantively relevant, 2) predictive of endogenous health behaviors, 3) exogenous—not influenced by other system variables, and 4) not directly related to BMI (and can be excluded from the BMI model). We based our causal model on substantive considerations. In particular, the selection of IVs was guided by published findings on the role of the built environment in health behaviors (30–34). We included an assessment of the strength of our full set of IVs to predict endogenous health behaviors with F tests of first-stage models. The set of IVs, noninstrument exogenous variables, and the endogenous health behaviors included in the model are further described in the Web material (Web Appendix and Web Table 1).

Table 2. Sex-Specific Descriptive Statistics for Study Participants Over the Course of the Study Period,[a] CARDIA Study, 1985–1986 to 2010–2011

Variable	Year of Study Period											
	0		7		10		15		20		25	
	No. or %	Mean (SD)	No. or %	Mean (SD)	No. or %	Mean (SD)	No. or %	Mean (SD)	No. or %	Mean (SD)	No. or %	Mean (SD)
Men												
Total no. of participants	2,327		1,836		1,755		1,619		1,535		1,517	
African-American race, %	49.7		45.3		45.9		43.8		42.1		43.1	
Current smoker, %	31.6		29.0		28.2		24.0		21.0		18.9	
Alcohol drinker, %	71.6		66.1		63.9		62.3		62.8		63.3	
Married, %	20.4		44.0		49.6		56.7		58.9		57.3	
Age, years		24.8 (3.6)		32.0 (3.6)		35.0 (3.6)		40.2 (3.6)		45.2 (3.5)		50.1 (3.6)
Education, years		13.8 (2.4)		14.7 (2.6)		14.9 (2.6)		15.2 (2.6)		15.4 (2.6)		15.5 (2.7)
Fast-food consumption, frequency/week		2.34 (2.51)		2.30 (2.72)		2.05 (2.17)		2.08 (2.53)		1.91 (2.28)		1.53 (2.31)
Physical activity, intensity units[b]												
Total physical activity[c]		5.22 (3.23)		4.30 (2.99)		4.23 (3.01)		4.25 (3.05)		4.07 (2.91)		4.07 (2.93)
Walking activity[d]		0.53 (0.56)		0.44 (0.51)		0.43 (0.52)		0.46 (0.52)		0.49 (0.54)		0.54 (0.54)
Nonwalking activity[d]		4.69 (3.10)		3.86 (2.89)		3.80 (2.87)		3.79 (2.92)		3.58 (2.76)		3.53 (2.77)
Combined walking, running, and biking[d]		1.58 (1.31)		1.35 (1.28)		1.32 (1.30)		1.37 (1.32)		1.37 (1.35)		1.46 (1.40)
All activities other than walking, running, and biking[d]		3.64 (2.49)		2.95 (2.27)		2.91 (2.31)		2.89 (2.27)		2.70 (2.14)		2.61 (2.11)
Women												
Total no. of participants	2,785		2,248		2,192		2,051		2,013		1,980	
African-American race, %	53.1		50.9		51.1		49.8		49.9		49.8	
Current smoker, %	29.3		25.1		23.2		20.4		17.9		15.3	
Alcohol drinker, %	52.2		46.2		44.2		44.2		47.2		48.5	
Married, %	23.8		43.6		47.8		51.0		52.6		50.6	
Age, years		24.9 (3.7)		32.1 (3.7)		35.0 (3.7)		40.2 (3.6)		45.2 (3.5)		50.1 (3.6)
Education, years		13.8 (2.1)		14.7 (2.3)		14.9 (2.4)		15.2 (2.6)		15.4 (2.6)		15.5 (2.7)
Fast-food consumption, frequency/week		1.72 (1.97)		1.66 (2.14)		1.54 (1.87)		2.08 (2.53)		1.91 (2.28)		1.53 (2.31)
Physical activity, intensity units[b]												
Total physical activity[c]		3.35 (2.51)		2.62 (2.21)		2.56 (2.24)		4.25 (3.05)		4.07 (2.91)		4.07 (2.93)
Walking activity[d]		0.57 (0.54)		0.49 (0.49)		0.50 (0.51)		0.46 (0.52)		0.49 (0.54)		0.54 (0.54)
Nonwalking activity[d]		2.78 (2.32)		2.13 (2.04)		2.06 (2.06)		3.79 (2.92)		3.58 (2.76)		3.53 (2.77)
Combined walking, running, and biking[d]		1.27 (1.17)		0.99 (1.05)		0.97 (1.09)		1.37 (1.32)		1.37 (1.35)		1.46 (1.4)
All activities other than walking, running, and biking[d]		2.08 (1.77)		1.63 (1.56)		1.58 (1.61)		2.89 (2.27)		2.70 (2.14)		2.61 (2.11)

Abbreviations: CARDIA, Coronary Artery Risk Development in Young Adults; SD, standard deviation.

[a] Data are shown for each of the 6 CARDIA examination periods included in the analysis, with year 0 being the study baseline.

[b] Physical activity intensity units have been rescaled by dividing by 100. Physical activity variables are shown for the 3 model specifications.

[c] Total physical activity intensity units equaled the sum of intensity units contributed by each of 13 separate activities assessed in CARDIA, including: running; biking; swimming; racket sports; exercise class or dancing; job activity such as lifting, carrying, or digging; home or leisure activity such as snow-shoveling, moving heavy objects, or weight-lifting; strenuous sports such as basketball, football, skating, or skiing; golf; calisthenics; home maintenance or gardening; and strenuous sports not otherwise listed (e.g., volleyball, table tennis).

[d] Intensity units for specific activities (e.g., walking vs. nonwalking activities) sum to total physical activity intensity units.

We compared our full-system-estimation IV-SEM with several other model-based approaches available in Stata software (StataCorp LP, College Station, Texas; see Table 1), including OLS regression with robust variance estimation (Stata's -regress- command), random-effects regression (-xtreg- command, "re" option), fixed-effects regression (-xtreg-, "fe" option), and single-equation-estimation IV-SEM (-ivregress-, "gmm" option). OLS and random-effects regression assume that there is no residual confounding or differential measurement error, although the random effects allow for individual variability in random errors. In contrast, differential unobserved heterogeneity is controlled in fixed-effects regression (time-invariant sources only) and single-equation-estimation IV-SEM (both time-invariant and time-varying sources). Fixed-effects regression can suffer from low statistical power due to loss of degrees of freedom (individual-level differencing) and a lack of within-person variability in the exposure of interest, and it does not allow estimation of time-invariant predictors.

Single-equation-estimation IV-SEM methods are less efficient than full-information IV-SEM methods. In addition, single-equation linear estimators (as in -ivregress-) are particularly susceptible to weak instruments; our method allows nonlinear estimation and has been shown to be robust in the presence of weak instruments (21, 22). We use 2 specifications of our full-system-estimation IV-SEM: 1) accounting for endogeneity due to time-invariant unobserved heterogeneity only and 2) accounting for time-varying unobserved heterogeneity as well. Our model is more flexible parametrically than other IV approaches, in terms of the availability of nonlinear functional forms for model equations and the use of the discrete factor method for modeling endogeneity due to unobserved heterogeneity. (The discrete factor method is further described in the Web Appendix.) In the present analysis, we assumed constant effect estimates over time and with respect to participant characteristics (no interaction).

We conducted analyses on 5,112 participants with data collected over 6 examination periods. Three participants were excluded from the original sample of 5,115 (1 dropped out of the study and 2 changed sex). Because of follow-up losses, there were 4,010 participants at year 7; 3,947 at year 10; 3,670 at year 15; 3,548 at year 20; and 3,497 at year 25 (23,858 observations in total). We tested the influence of loss to follow-up with sensitivity analysis using inverse probability weighting by examination participation; these models are not presented, as results did not differ from those of unadjusted models. There were complete data on community-level indicators and participant age, sex, and race. We used regression prediction models to fill in missing individual-level data (described in the Web Appendix and Web Table 2). We used an α level of 0.05 for statistical significance. We used Stata (version 13) and Fortran (Intel Fortran Compiler; Intel Corporation, Santa Clara, California) for all analyses.

RESULTS

Over the 25-year study period, smoking and fast-food consumption declined among both men and women, and

Table 3. Fit Statistics for Each of the 3 Sex-Specific Physical Activity Specifications of the Instrumental-Variables Simultaneous Equations Model, CARDIA Study, 1985–1986 to 2010–2011

Variable	Total Physical Activity Model				Walking Model				Walking, Running, and Biking Model			
	Men		Women		Men		Women		Men		Women	
	F Test[a]	*P* Value	*F* Test	*P* Value	*F* Test	*P* Value	*F* Test	*P* Value	*F* Test	*P* Value	*F* Test	*P* Value
Current smoking	1.2	0.19	1.5	0.04	1.2	0.19	1.5	0.04	1.2	0.19	1.5	0.04
Alcohol consumption	3.6	<0.01	2.2	<0.01	3.6	<0.01	2.2	<0.01	3.6	<0.01	2.2	<0.01
Marital status	12.1	<0.01	2.6	<0.01	12.1	<0.01	2.6	<0.01	12.1	<0.01	2.6	<0.01
Fast-food consumption	3.8	<0.01	1.9	<0.01	3.8	<0.01	1.9	<0.01	3.8	<0.01	1.9	0.03
Physical activity[b]												
Total physical activity	1.6	0.02	2.6	<0.01								
Walking activity					4.3	<0.01	3.3	<0.01				
Nonwalking activity					1.6	0.02	2.2	<0.01				
Walking, running, and biking									2.5	<0.01	2.3	<0.01
Activities other than walking, running, and biking									2.0	<0.01	2.4	<0.01

Abbreviation: CARDIA, Coronary Artery Risk Development in Young Adults.

[a] *F* tests of the first-stage models provide evidence of the combined strength of all instrumental variables to predict the endogenous health behavior outcome. Rejection of the *F* test ($P < 0.05$) indicates that our set of instruments provides good identification for that endogenous health behavior variable. The set of instrumental (or other exogenous) variables is described in the text, and they comprise indicators of community prices (e.g., housing, food), street connectivity, neighborhood food resources and physical activity facilities, and Census tract-level socioeconomic status (income, education).

[b] Three physical activity specifications were fitted: total physical activity; walking/nonwalking; and walking, running, biking/all other activities. For each specification, 4 other health behaviors (hypothesized to be associated with body mass index) were modeled, including current smoking, alcohol consumption, marital status, and fast-food consumption.

engagement in physical activity declined among men, with the exception of walking, which remained stable (Table 2). Changes in physical activity were less clear among women.

Table 3 presents results from sex-specific model F tests for each of the 3 physical activity specifications, reflecting the strength of the full set of IVs in predicting endogenous health behaviors. F test statistics were small and well below recommended thresholds (e.g., ≥10) (35) for considering the models strongly identified; that is, our models suffered from "weak" instruments. Even in the presence of weak instruments, P values for first-stage model F tests were statistically significant for all models except smoking among men.

With the exception of generalized method of moments (GMM) IV-SEM, model estimates were surprisingly similar across the 6 modeling approaches (Table 4). The similarity of point estimates from approaches that correct (fixed effects and full-information maximum likelihood (FIML) IV-SEM) and do not correct (OLS and random effects) for endogeneity due to unobserved heterogeneity supports a lack of residual confounding or differential measurement error in multivariable-adjusted models, which is contrary to our expectation and unlikely to hold in general. Of the 2 FIML IV-SEMs, those that accounted for time-invariant unobserved heterogeneity only generally had a higher value for the log-likelihood function.

Point estimates from the GMM IV-SEM were frequently substantively different from those of other models, as well as much less precise (Table 4). As an example, in both men and women, all models supported a negative association between smoking and BMI, with the exception of the GMM IV-SEM, which yielded a positive (albeit statistically nonsignificant) estimate. Current smokers had a lower BMI than nonsmokers (men: −0.88 (95% confidence interval (CI): −1.49, −0.27); women: −0.61 (95% CI: −1.18, −0.05)), based on the FIML IV-SEM accounting for time-invariant unobserved heterogeneity (10th column of Table 4). Among men, some GMM IV-SEM estimates were consistent with models that did not account for unobserved heterogeneity; for example, there was an apparent negative association between alcohol consumption and BMI based on OLS, random effects, and GMM IV-SEM, which was not supported in fixed-effects regression and FIML IV-SEM.

Aside from GMM IV-SEM, model estimates were generally consistent with expectation in both men and women (Table 4). Physical activity and smoking were negatively associated with BMI, while fast-food consumption and marriage were positively associated with BMI. Among women, but less consistently among men, alcohol consumption was negatively associated with BMI. Although we did not formally examine effect measure modification by sex, the variables alcohol consumption, physical activity, and fast-food consumption appeared to be more strongly predictive of BMI in women than in men. For example, among women, the BMI differences associated with a 100-unit increase in total physical activity were −0.18 (95% CI: −0.23, −0.12) and −0.27 (95% CI: −0.39, −0.16) in the FIML IV-SEM models accounting for, respectively, time-invariant unobserved heterogeneity and time-invariant and time-varying unobserved heterogeneity, as compared with −0.09 (95% CI: −0.16, −0.02) and −0.09 (95% CI: −0.17, −0.01), respectively, among men.

DISCUSSION

We have presented results from a semiparametric full-system-estimation approach to IV-SEM analysis (FIML IV-SEM) of BMI over 25 years of follow-up in the CARDIA cohort. Using a system of equations to predict BMI as a function of weight-related behaviors, we corrected for residual confounding or differential measurement error (endogeneity due to unobserved heterogeneity). Substantively, our findings were consistent with expectation, with BMI being negatively associated with physical activity and smoking and positively associated with fast-food consumption. In addition, marriage was positively associated with BMI, and alcohol consumption was negatively associated with BMI. Estimates from models that adjusted and did not adjust for unobserved heterogeneity were generally similar, indicating a relative lack of endogeneity after controlling for observed covariates, with the exception of OLS regression, which tended to overstate the significance of health behaviors in influencing BMI. Effect estimates from a single-equation-estimation approach to IV-SEM analysis (GMM IV-SEM) were notably different from other estimates, reflecting the challenge of using standard linear IV methods in the presence of weak instruments.

We hypothesized greater differences across modeling approaches, and in particular we did not anticipate marked similarity between estimates that did not account for residual confounding and differential measurement error (OLS and random effects) and those that did (fixed effects and FIML IV-SEM). The largest differences were observed for GMM IV-SEM regression estimates, which were also the least precise. The differences between GMM and FIML IV-SEM approaches are especially noteworthy, as GMM IV-SEM is available in standard statistical software and more accessible to researchers. Differences between GMM and FIML IV-SEM estimates probably reflect FIML's allowance of nonlinear estimators and our use of the discrete factor method (21, 22). GMM IV-SEM is a single-equation-estimation approach (fitting within the 2-stage least squares framework) and lacks the efficiency of full-system estimators, which not only account for unmeasured confounding and differential measurement error (unobserved heterogeneity) but use that information for more precise parameter estimation.

IV methods are known to be sensitive to weak identification. Our comparison illustrates the potential for severe bias of IV approaches in the absence of strong instruments. In contrast, the consistency of results from our full-system-estimation IV-SEM and other models, particularly fixed-effects regression, is illustrative of the robust nature of our approach in the presence of weak instruments. As was shown in previous work, our FIML IV-SEM has substantially stronger estimation performance in the presence of weak instruments (21, 22). Our results illustrate that, in the presence of weak instruments, standard IV approaches can be less preferable than a non-IV method, such as random- or fixed-effects regression.

In practice, several considerations are likely to drive the decision about which modeling approach to adopt, as outlined in Table 1. Assuming that unmeasured confounding or differential measurement error is considered a threat to validity, fixed-effects models or IV methods can be used to account for such nonrandom unobserved heterogeneity. Fixed-effects

Table 4. Regression Coefficients (β) for the Effects of Physical Activity and Other Health Behaviors on Body Mass Index,[a] CARDIA Study, 1985–1986 to 2010–2011[b]

Model and Variable	OLS, Robust Variance[c]		Repeated-Measures Regression, Random Effects[d]		Fixed-Effects Regression[e]		GMM IV-SEM[f]		FIML IV-SEM[g] (Time-Invariant Residual Confounding)		FIML IV-SEM[h] (Time-Invariant and -Varying Residual Confounding)	
	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI
Men												
Model 1[i]												
Total physical activity, intensity units	−0.13	−0.18, −0.08	−0.08	−0.10, −0.06	−0.08	−0.10, −0.06	0.12	−0.33, 0.57	−0.09	−0.16, −0.02	−0.09	−0.17, −0.01
Current smoking, yes/no	−1.24	−1.66, −0.83	−0.86	−1.03, −0.68	−0.78	−0.97, −0.60	0.99	−3.01, 4.99	−0.88	−1.49, −0.27	−0.98	−1.37, −0.59
Alcohol consumption, yes/no	−0.93	−1.29, −0.58	−0.17	−0.30, −0.04	−0.10	−0.23, 0.03	−5.16	−7.58, −2.73	−0.12	−0.39, 0.15	−0.03	−0.27, 0.20
Married, yes/no	0.55	0.19, 0.91	0.19	0.05, 0.32	0.13	−0.01, 0.27	1.30	0.05, 2.54	0.46	0.11, 0.81	0.55	0.16, 0.94
Fast-food consumption, frequency/week	0.13	0.06, 0.19	0.05	0.03, 0.07	0.04	0.02, 0.06	0.23	−0.23, 0.68	0.05	0.01, 0.09	0.08	0.01, 0.14
Model 2[j]												
Walking, intensity units	−0.10	−0.33, 0.14	−0.13	−0.23, −0.04	−0.13	−0.23, −0.03	0.45	−1.55, 2.45	−0.10	−0.31, 0.12	−0.11	−0.31, 0.09
Nonwalking, intensity units	−0.13	−0.18, −0.08	−0.08	−0.10, −0.06	−0.08	−0.10, −0.05	0.08	−0.43, 0.59	−0.09	−0.13, −0.04	−0.04	−0.09, 0.01
Model 3[k]												
Walking, running, and cycling, intensity units	−0.42	−0.53, −0.32	−0.18	−0.22, −0.13	−0.15	−0.20, −0.11	−0.92	−1.97, 0.13	−0.15	−0.25, −0.05	−0.16	−0.30, −0.02
All other physical activity, intensity units	0.01	−0.06, 0.07	−0.04	−0.07, −0.01	−0.04	−0.07, −0.02	0.78	0.10, 1.46	−0.09	−0.15, −0.03	0.08	0.01, 0.15
Women												
Model 1												
Total physical activity, intensity units	−0.33	−0.40, −0.25	−0.22	−0.25, −0.19	−0.21	−0.24, −0.17	−0.30	−0.97, 0.37	−0.18	−0.23, −0.12	−0.27	−0.39, −0.16
Current smoking, yes/no	−0.65	−1.16, −0.13	−0.95	−1.17, −0.73	−1.02	−1.25, −0.78	2.94	−2.45, 8.33	−0.61	−1.18, −0.05	−0.76	−1.38, −0.15
Alcohol consumption, yes/no	−1.33	−1.72, −0.94	−0.45	−0.60, −0.30	−0.36	−0.52, −0.20	0.27	−3.18, 3.73	−0.36	−0.71, −0.02	−0.54	−0.92, −0.16
Married, yes/no	0.20	−0.27, 0.67	0.32	0.15, 0.49	0.35	0.17, 0.52	−3.67	−7.64, 0.30	0.37	0.06, 0.68	0.39	−0.01, 0.80
Fast-food consumption, frequency/week	0.25	0.16, 0.35	0.12	0.08, 0.15	0.11	0.07, 0.14	−0.07	−1.15, 1.00	0.09	0.02, 0.15	0.14	0.06, 0.21
Model 2												
Walking, intensity units	−0.31	−0.59, −0.02	−0.38	−0.50, −0.25	−0.38	−0.51, −0.25	−1.95	−5.82, 1.93	−0.28	−0.51, −0.06	−0.16	−0.38, 0.06
Nonwalking, intensity units	−0.33	−0.40, −0.25	−0.20	−0.24, −0.17	−0.19	−0.22, −0.15	0.03	−1.02, 1.09	−0.18	−0.23, −0.13	−0.13	−0.20, −0.06

Table continues

Table 4. Continued

Model and Variable	OLS, Robust Variance[c]		Repeated-Measures Regression, Random Effects[d]		Fixed-Effects Regression[e]		GMM IV-SEM[f]		FIML IV-SEM[g] (Time-Invariant Residual Confounding)		FIML IV-SEM[h] (Time-Invariant and -Varying Residual Confounding)	
	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI
Model 3												
Walking, running, and cycling, intensity units	−0.70	−0.83, −0.58	−0.34	−0.40, −0.27	−0.30	−0.37, −0.24	−0.19	−2.10, 1.72	−0.38	−0.50, −0.27	−0.40	−0.51, −0.29
All other physical activity, intensity units	−0.12	−0.22, −0.02	−0.16	−0.20, −0.11	−0.15	−0.20, −0.11	−0.37	−1.70, 0.97	−0.05	−0.13, 0.04	−0.06	−0.15, 0.03

Abbreviations: BMI, body mass index; CARDIA, Coronary Artery Risk Development in Young Adults; CI, confidence interval; FIML, full-information maximum likelihood; GMM, generalized method of moments; OLS, ordinary least squares; IV-SEM, instrumental-variables simultaneous equations model.

[a] Weight (kg)/height (m)^2.

[b] Regression coefficients were obtained from OLS, repeated-measures random-effects regression, fixed-effects regression, and 3 instrumental-variable models. Multivariable models adjusted for age, age squared, sex, study center, participant educational attainment, participant income, and educational attainment of the participant's parents. The model for women additionally adjusted for number of children.

[c] Multivariable-adjusted OLS regression using Stata's -`regress`- command (StataCorp LP, College Station, Texas) with the robust variance option.

[d] Repeated-measures random-effects regression using Stata's -`xtreg`- command with the "`re`" option.

[e] Repeated-measures fixed-effects regression using Stata's -`xtreg`- command with the "`fe`" option.

[f] IV-SEM using Stata's -`ivregress`- command with the "`gmm`" option.

[g] IV-SEM in Fortran (Intel Corporation, Santa Clara, California), accounting for unobserved time-invariant residual confounding.

[h] IV-SEM in Fortran, accounting for unobserved time-invariant and time-varying residual confounding.

[i] Model 1 included a single equation for CARDIA physical activity intensity units, rescaled by dividing by 100.

[j] Model 2 included 2 equations for physical activity: walking and nonwalking. Walking and nonwalking components of total physical activity were derived by summing intensity units for walking and intensity units for all 12 nonwalking components, respectively (queried about in the CARDIA survey). In addition to walking and nonwalking physical activity, the model system included equations for other endogenous BMI predictors included in the total physical activity model: smoking, alcohol consumption, marital status, and fast-food consumption. Beta coefficients for other BMI predictors were similar to those of the total physical activity model and are not displayed.

[k] Model 3 included 2 equations for physical activity: 1) walking, running, and biking; and 2) all other forms of physical activity. Distinct components of total physical activity were derived by summing intensity units for walking, running, and biking and intensity units for all 10 other components, respectively (queried about in the CARDIA survey). In addition to equations for walking/running/biking and all other physical activity, the model system included equations for other endogenous BMI predictors included in the total physical activity model: smoking, alcohol consumption, marital status, and fast-food consumption. Beta coefficients for other BMI predictors were similar to those of the total physical activity model and are not displayed.

models are limited to studies with more than 1 observation period, relatively large samples, and estimation of effects for exposures that change sufficiently over the observation period; in the absence of one or more of these features, IV methods can be considered. IV approaches require that valid instruments can be identified. In the presence of strong instruments, IV methods are generally robust to many parametric assumptions; in the presence of weak instruments, more robust methods (such as the FIML IV-SEM) may be necessary. Furthermore, as in any modeling exercise, results may be more or less sensitive to specification assumptions relating to the linearity of outcome variables or multivariate normality of random effects, and approaches that limit these assumptions will be preferable.

Our FIML IV-SEM estimates that accounted only for time-invariant unobserved heterogeneity were generally similar to those that accounted for both time-invariant and time-varying unobserved heterogeneity. Although it is conceptually preferable to account for time-varying unobserved heterogeneity, for several reasons we prefer the models that accounted for time-invariant unobserved heterogeneity only. First, when estimates differed meaningfully (such as all other physical activity in the walking/running/biking model among men), the results from the model that accounted for time-invariant unobserved heterogeneity only were more consistent with substantive expectation (in the exception noted, physical activity was inversely associated with BMI). In addition, the models with time-invariant unobserved heterogeneity accommodated an equal or greater number of levels for unobserved heterogeneity, which we would expect to improve our control for endogeneity. Finally, results from the time-invariant-only models were generally more precise.

Substantively, our results confirm the importance of physical activity, smoking, and fast-food consumption in shaping body mass (36–38). The association between alcohol consumption and BMI has been equivocal in epidemiologic studies (39). In our heterogeneity-corrected analysis, alcohol consumption was negatively associated with BMI among women, as supported by prior work (40, 41), but not among men, suggesting possible sex differences. Our findings and those of others (42, 43) indicate that marriage is positively associated with BMI. The roles of alcohol consumption and marriage in BMI merit further study.

Although our method accounted for confounding due to omitted variables and differential measurement error, it may have residual bias due to residential selection. Community-level indicators served as IVs; however, the potential for informative residential selection is a recognized challenge in studies of neighborhood exposures and health (44, 45). We considered including community dummy variables as model covariates, but this proved infeasible due to the very large number of communities with small numbers of participants. Instead, we adjusted for baseline study center, hypothesizing that participant residence was independent of error components conditional on baseline study center. The similarity of estimates derived from our model and from fixed-effects regression indicates that selective migration due to unobserved time-invariant individual characteristics was not an appreciable source of bias.

The use of community-level indicators as IVs is common in the econometrics and epidemiologic literature but may have contributed to the weak identification. Only 1 community indicator, cigarette price, was directly related to cigarette smoking, which was the most poorly identified health behavior. IVs are assumed to be exogenous and, ideally, are strongly predictive of endogenous predictors but not directly predictive of the system outcome (BMI). Achieving this balance is the greatest challenge in IV analysis, and one's success in doing so cannot be directly tested. A major contribution of our method is the ability to obtain consistent estimates even in the presence of weak instruments, which allows us to focus on IVs less likely to be within the model system, such as community indicators. We note that weak identification may also reflect nondifferential measurement errors in our IVs.

Our model does not eliminate the possibility of bias, such as bias from model misspecification, but the robust findings across multiple regression approaches (aside from GMM IV-SEM) are supportive of causal estimates. In addition to our statistical modeling approach, a strength of our study was our extensive set of candidates for IVs from a comprehensive set of community-level data. Our finding of multiple levels of time-invariant and time-varying unobserved heterogeneity is a testament to our rich data set and highlights our ability to capture significant unobserved heterogeneity.

In conclusion, we used 25 years of CARDIA data to jointly model health behaviors and BMI. Our analysis yielded consistent estimates of the influences of diet, physical activity, smoking, and other variables on BMI, and it accounted for endogeneity due to unobserved heterogeneity stemming, we hypothesized, from residual confounding and differential measurement error. Our results confirm the importance of physical inactivity and fast-food consumption in relation to weight gain, and they support the development of policy and intervention efforts in these areas. In addition, our findings indicate that marital status and alcohol consumption may play underappreciated roles in body mass. The large differences between our full-system-estimation IV-SEM, as well as fixed- and random-effects regression, and the single-equation-estimation method that used standard statistical software (Stata) illustrate that, in the presence of weak instruments, a non-IV approach may be preferable. The potential for large bias when using simple IV models in the presence of weak instruments is added support for our method—for which estimates were robust across estimation approaches. This work contributes to the growing body of literature related to the use of IV methods in epidemiologic practice.

ACKNOWLEDGMENTS

Author affiliations: Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Katie A. Meyer, Barry M. Popkin, Penny Gordon-Larsen); Department of Economics, College of Arts and Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (David K. Guilkey); Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (David K. Guilkey, Hsiao-Chuen Tien, Barry M. Popkin, Penny Gordon-Larsen); and Department of Quantitative

Health Sciences, University of Massachusetts Medical School, Worcester, Massachusetts (Catarina I. Kiefe).

This work was funded by the National Heart, Lung, and Blood Institute (NHLBI) (grant R01HL104580). The Coronary Artery Risk Development in Young Adults (CARDIA) Study is supported by contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN268200900041C from the NHLBI, the Intramural Research Program of the National Institute on Aging (NIA), and an intraagency agreement between the NIA and the NHLBI (grant AG0005). For general support, we are grateful to the Carolina Population Center, University of North Carolina at Chapel Hill (grant R24HD050924 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development), the Nutrition Obesity Research Center, University of North Carolina (grant P30DK56350 from the National Institute for Diabetes and Digestive and Kidney Diseases), and the UNC Center for Environmental Health and Susceptibility, University of North Carolina (grant P30ES010126 from the National Institute of Environmental Health Sciences).

The National Institutes of Health had no role in the design or conduct of the study; the collection, management, analysis, or interpretation of the data; or the preparation, review, or approval of the manuscript.

Conflict of interest: none declared.

REFERENCES

1. Butland B, Jebb S, Kopelman P, et al. *Tackling Obesities: Future Choices*. (Foresight Project report). London, United Kingdom: Government Office for Science; 2007.
2. Huang TT, Drewnosksi A, Kumanyika S, et al. A systems-oriented multilevel framework for addressing obesity in the 21st century. *Prev Chronic Dis*. 2009;6(3):A82.
3. Briscoe J, Akin J, Guilkey D. People are not passive acceptors of threats to health: endogeneity and its consequences. *Int J Epidemiol*. 1990;19(1):147–153.
4. Chou SY, Grossman M, Saffer H. An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *J Health Econ*. 2004;23(3):565–587.
5. Rashad I. Structural estimation of caloric intake, exercise, smoking, and obesity. *Q Rev Econ Finance*. 2006;46(2): 268–283.
6. French MT, Norton EC, Fang H, et al. Alcohol consumption and body weight. *Health Econ*. 2010;19(7):814–832.
7. Ng SW, Norton EC, Guilkey DK, et al. Estimation of a dynamic model of weight. *Empir Econ*. 2012;42(2):413–443.
8. Cebu Study Team. Underlying and proximate determinants of child health: the Cebu Longitudinal Health and Nutrition Study. *Am J Epidemiol*. 1991;133(2):185–201.
9. Subar AF, Kipnis V, Troiano RP, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol*. 2003;158(1):1–13.
10. Neuhouser ML, Di C, Tinker LF, et al. Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. *Am J Epidemiol*. 2013;177(6): 576–585.
11. Greene WH. *Econometric Analysis*. 7th ed. Upper Saddle River, NJ: Prentice Hall; 2011.
12. Guilkey DK, Hutchinson PL. Overcoming methodological challenges in evaluating health communication campaigns: evidence from rural Bangladesh. *Stud Fam Plann*. 2011;42(2): 93–106.
13. Rindfuss RR, Guilkey DK, Morgan SP, et al. Child-care availability and fertility in Norway. *Popul Dev Rev*. 2010;36(4): 725–748.
14. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000;29(4):722–729.
15. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology*. 2006;17(4):360–372.
16. Burgess S, Granell R, Palmer TM, et al. Lack of identification in semiparametric instrumental variable models with binary outcomes. *Am J Epidemiol*. 2014;180(1):111–119.
17. Palmer TM, Sterne JA, Harbord RM, et al. Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol*. 2011; 173(12):1392–1403.
18. Rassen JA, Schneeweiss S, Glynn RJ, et al. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. *Am J Epidemiol*. 2009;169(3): 273–284.
19. Trompet S, Jukema JW, Katan MB, et al. Apolipoprotein E genotype, plasma cholesterol, and cancer: a Mendelian randomization study. *Am J Epidemiol*. 2009;170(11): 1415–1421.
20. Au Yeung SL, Jiang CQ, Cheng KK, et al. Evaluation of moderate alcohol use and cognitive function among men using a Mendelian randomization design in the Guangzhou Biobank Cohort Study. *Am J Epidemiol*. 2012;175(10):1021–1028.
21. Mroz TA. Discrete factor approximations in simultaneous equation models: estimating the impact of a dummy endogenous variable on a continuous outcome. *J Econom*. 1999;92(2):233–274.
22. Guilkey DK, Lance PM. Program impact estimation with binary outcome variables: Monte Carlo results for alternative estimators and empirical examples. In: Sickles RC, Horrace WC, eds. *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications*. New York, NY: Springer Publishing Company; 2014:5–46.
23. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc*. 1996; 91(434):444–455.
24. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41(11):1105–1116.
25. Jacobs DR, Hahn LP, Haskell WL, et al. Validity and reliability of short physical activity history: CARDIA and the Minnesota Heart Health Program. *J Cardiopulm Rehabil*. 1989;9(11): 448–459.
26. Liu K, Slattery M, Jacobs D Jr, et al. A study of the reliability and comparative validity of the CARDIA dietary history. *Ethn Dis*. 1994;4(1):15–27.
27. Heckman J, Singer B. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*. 1984;52(2):271–320.
28. Angeles G, Guilkey DK, Mroz TA. Purposive program placement and the estimation of family planning program effects in Tanzania. *J Am Stat Assoc*. 1998;93(443):884–899.
29. Laird N. Nonparametric maximum likelihood estimation of a mixing distribution. *J Am Stat Assoc*. 1978;73(364):805–811.
30. Morland K, Wing S, Diez Roux A. The contextual effect of the local food environment on residents' diets: the Atherosclerosis Risk in Communities Study. *Am J Public Health*. 2002;92(11): 1761–1767.

31. Meyer KA, Guilkey DK, Ng SW, et al. Sociodemographic differences in fast food price sensitivity. *JAMA Intern Med*. 2014;174(3):434–442.
32. Hou N, Popkin BM, Jacobs DR Jr, et al. Longitudinal associations between neighborhood-level street network with walking, bicycling, and jogging: the CARDIA study. *Health Place*. 2010;16(6):1206–1215.
33. Boone-Heinonen J, Gordon-Larsen P, Kiefe CI, et al. Fast food restaurants and food stores: longitudinal associations with diet in young to middle-aged adults: the CARDIA study. *Arch Intern Med*. 2011;171(13):1162–1170.
34. Andreyeva T, Long MW, Brownell KD. The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. *Am J Public Health*. 2010; 100(2):216–222.
35. Stock JH, Wright JH, Yogo M. A survey of weak instruments and weak identification in generalized method of moments. *J Bus Econ Stat*. 2002;20(4):518–529.
36. Hankinson AL, Daviglus ML, Bouchard C, et al. Maintaining a high physical activity level over 20 years and weight gain. *JAMA*. 2010;304(23):2603–2610.
37. Lewis CE, Smith DE, Wallace DD, et al. Seven-year trends in body weight and associations with lifestyle and behavioral characteristics in black and white young adults: the CARDIA study. *Am J Public Health*. 1997;87(4): 635–642.
38. Pereira MA, Kartashov AI, Ebbeling CB, et al. Fast-food habits, weight gain, and insulin resistance (the CARDIA study): 15-year prospective analysis. *Lancet*. 2005;365(9453):36–42.
39. Suter PM. Is alcohol consumption a risk factor for weight gain and obesity? *Crit Rev Clin Lab Sci*. 2005;42(3):197–227.
40. Wang L, Lee IM, Manson JE, et al. Alcohol consumption, weight gain, and risk of becoming overweight in middle-aged and older women. *Arch Intern Med*. 2010;170(5):453–461.
41. Thomson CA, Wertheim BC, Hingle M, et al. Alcohol consumption and body weight change in postmenopausal women: results from the Women's Health Initiative. *Int J Obes (Lond)*. 2012;36(9):1158–1164.
42. Sobal J, Rauschenbach B, Frongillo EA. Marital status changes and body weight changes: a US longitudinal analysis. *Soc Sci Med*. 2003;56(7):1543–1555.
43. French SA, Jeffery RW, Forster JL, et al. Predictors of weight change over two years among a population of working adults: the Healthy Worker Project. *Int J Obes Relat Metab Disord*. 1994;18(3):145–154.
44. Boone-Heinonen J, Gordon-Larsen P, Guilkey DK, et al. Environment and physical activity dynamics: the role of residential self-selection. *Psychol Sport Exerc*. 2011;12(1): 54–60.
45. Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med*. 2004;58(10):1929–1952.