Empir Econ
DOI 10.1007/s00181-010-0403-6

Body mass index in families: spousal correlation, endogeneity, and intergenerational transmission

Jason Abrevaya · Hongfei Tang

Received: 26 February 2008 / Accepted: 2 June 2010

Abstract Previous studies have documented spousal and intergenerational correlations in body mass index (BMI) but few have considered familial weight data augmented with socioeconomic and behavioral control variables. This article considers a U.S. dataset that contains such information on husbands, wives, and grown children. Although certain variables (like education, race, and smoking status) are helpful in explaining an individual's BMI, the BMI of one's spouse (or parents) remains the most significant predictor of BMI. To help distinguish between correlation and causality in the married-adult BMI regressions, we consider two alternative approaches for dealing with possible endogeneity (due to omitted variables): (1) including spousal variables to proxy for omitted variables and (2) modeling spousal BMI in a hierarchical framework to explicitly allow for a "couple" effect. The results suggest endogeneity of educational attainment, but not smoking status, and support prior research that finds different associations of BMI with income for husbands and wives. For grown children, parental BMI and smoking status are identified as significant predictors.

Keywords Body mass index · Endogeneity · Familial correlation

JEL Classification I10 · I12

J. Abrevaya (✉)
Department of Economics, The University of Texas at Austin, Austin, TX 78712, USA
e-mail: abrevaya@eco.utexas.edu

H. Tang
Stillman School of Business, Seton Hall University, South Orange, NJ 07079, USA
e-mail: hongfei.tang@shu.edu

Published online: 09 October 2010

1 Introduction

The prevalence of overweight and obesity has increased dramatically in the United States over the past two decades. Body mass index (BMI), defined as weight/height^2 (in units kg/m^2), is commonly used to classify *overweight* (BMI ≥ 25) and *obesity* (BMI ≥ 30) among adults. Figure 1, using estimates from Flegal et al. (2002), depicts the rise in obesity prevalence in the United States between 1960 and 2000. Whereas the prevalence of obesity remained fairly unchanged between 1960 and 1980, it has dramatically increased since 1980, with over 30% of American adults currently classified as obese (and nearly two-thirds classified as overweight).

There are a variety of elevated health risks associated with overweight and obesity, including hypertension, Type 2 diabetes, heart disease, stroke, and some forms of cancers (see, for example, Manson et al. (2004)). Finkelstein et al. (2004) estimate that obesity-attributable medical expenditures in the United States total roughly $80 billion annually and account for 5.7% of total adult medical expenditures.[1] Due to the tremendous costs associated with overweight and obesity, reducing their prevalence has become a top public-health priority. One complicating factor, however, is that predisposition to weight gain has a significant genetic component. Public policy can only effectively target weight issues to the extent that obesity and overweight are also determined by behavioral and/or environmental factors. (Such a concern is far less important for policies targeting other "risky" behaviors like smoking and drinking.) Although many medical studies have pointed to genetic factors as the key determinants of adult weight, the dramatic recent increase in overweight and obesity suggests that other factors must be increasing in importance.

Several recent studies have argued that agricultural and technological innovations are at least partially responsible for the recent increase in overweight and obesity. Philipson and Posner (2003) and Lakdawalla and Philipson (2009) argue that such innovations have lowered the cost of consuming calories and raised the cost of expending calories. Cutler et al. (2003) also point to increased caloric intake (through mass-produced-food consumption) and note that the biggest change in weight profile has occurred for married (and working) women, the group most able to take advantage of the technological advances. Among children, Wilson (2006) suggests that the dramatic increase in obesity has been caused by a shift in time allocation toward indolent activities (such as watching television and playing video games). The conclusions from these studies are echoed in a recent report by the World Health Organization (2003):

> The rising epidemic reflects the profound changes in society and in behavioural patterns of communities over recent decades. While genes are important in determining a person's susceptibility to weight gain, energy balance is determined by calorie intake and physical activity. Thus societal changes and worldwide nutrition transition are driving the obesity epidemic. Economic growth, mod-

[1] In a prior study, Wolf and Colditz (1998) had estimated that the total cost attributable to obesity in 1995 was $99.2 billion in 1995, with $51.6 billion for direct medical costs and the rest for indirect costs due to lost productivity.

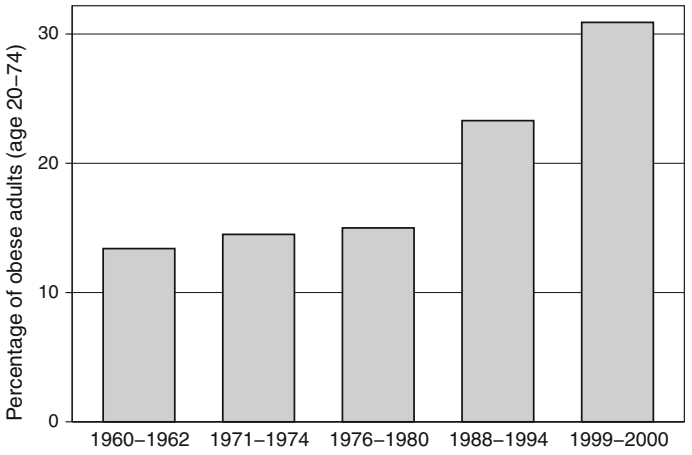

Fig. 1 Trends in obesity prevalence

ernization, urbanization and globalization of food markets are just some of the forces thought to underlie the epidemic.

Given the apparent increase in the importance of non-genetic factors, this study aims to develop a better understanding of familial relationships in body weight while controlling for important socioeconomic variables. Figure 2 summarizes the sources of familial correlations in BMI. Several research studies in the natural sciences have documented the spousal correlation in BMI (see, for example, Jeffery and Rick (2002) and the cites therein), with (simple) correlations generally in the range between 0.1 and 0.2. The spousal correlation arises from two primary sources, assortative matching (in choosing one's spouse) and common environment (including dietary habits, exercise routines, distance to restaurants, income, etc.). Parent–child correlations in BMI also arise from two primary sources, genetics and common environment. Several studies (including Sacerdote (2004), Stunkard et al. (1986), and Vogler et al. (1995)) have used data on adopted children to document the importance of the genetic component.[2] However, as pointed out above, the increasing trend in overweight and obesity among children suggests an increasing importance of the environmental component. As highlighted by some recent research, controlling for observables that might influence BMI is crucial for examining this environmental component. For example, Anderson et al. (2003) find that a child is more likely to be overweight if his/her mother works more hours, even after controlling for mother's weight status. Relatedly, Taveras et al. (2005) find that the frequency of eating family dinners is inversely related to overweight prevalence.

Given the importance of increasing BMI as a public-health concern, a better understanding of how social and family variables influence weight is critical for public

[2] For example, the parent–child BMI correlation is found to be much larger with biological parents than with adoptive parents. Studies on twins (see, e.g., Maes et al. (1997)) have arrived at similar conclusions by comparing the correlations of monozygotic twins to dizygotic twins (and other siblings).

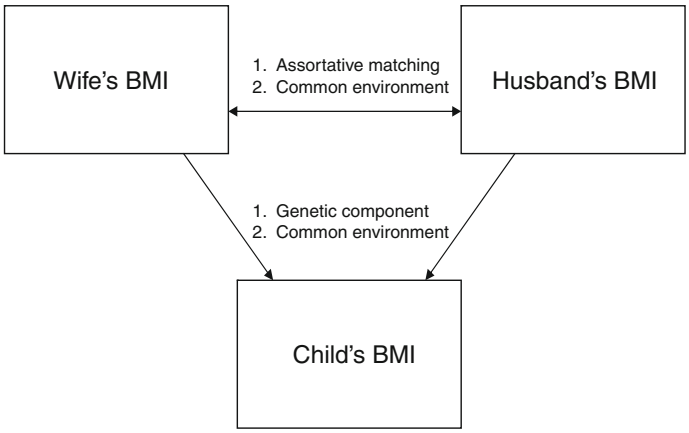

Fig. 2 Familial BMI correlations

policy issues. In this article, we utilize a large cross-sectional dataset (the Community Tracking Study) in the United States with information on husbands, wives, and grown children in order to examine the familial relationships in weight status and the determinants of weight status. The dataset contains a wide range of demographic variables (including age, race, and education), economic variables (including employment status and income), and behavioral variables (including smoking status). By using a comprehensive set of socioeconomic controls, this study builds upon other research that has documented familial correlations in BMI and other weight-related measures (for example, Mamum et al. (2005), Price et al. (2000), and Rice et al. (1997)). The examination of the parent–child relationship with control variables has been considered recently by Anderson et al. (2003) and Classen and Hokayem (2005). Both studies use the National Longitudinal Survey of Youth (NLSY) and find that mother's obesity status and education are significant predictors of childhood obesity. Unfortunately, the NLSY does not contain any information about fathers. One study that does utilize data on both spouses is Wilson (2002), which considers spousal correlations in health status using the Health and Retirement Survey (HRS). However, the focus of Wilson (2002) is upon general measures of health status (rather than weight status specifically), and the data source restricts the analysis to an older subpopulation of married individuals.

The initial focus is upon the spousal relationship, where the availability of data for *both* husbands and wives allows us to deal with possible endogeneity (due to omitted variables). In particular, we consider two alternative approaches to deal with endogeneity: (1) including spousal variables to proxy for omitted variables and (2) modeling spousal BMI in a hierarchical framework to explicitly allow for a "couple effect." The first approach can be viewed as a *predictive* regression in which own and spousal characteristics are used to predict own BMI, whereas the second approach is a more structural regression in which the couple effect (correlated with own and spousal characteristics) enters into own BMI. Although we recognize that causal interpretations should still be made cautiously here (due to the lack of convincing instrumental variables), the hierarchical nature of the data allows us to improve markedly upon a regression analysis without spousal information.

The second focus of this article is upon the BMI outcomes for grown children (ages 18–30). Although the Community Tracking Study does not provide weight and height information for children under the age of 18, it does provide data on adult children (ages 18–30) still living with their parents. This unique feature of the data allows us to include parental control variables (such as BMI and smoking status) in the grown child's BMI regression. As far as we know, this study is the first to conduct such a predictive BMI regression for grown children in the presence of a wide range of own and parental control variables.

The outline of this article is as follows. Section 2 describes the Community Tracking Study data that are used for the empirical analysis. Taking advantage of the fact that the data contain information on all adults within a household and some children, two different samples are considered: (i) a sample of married adults and (ii) a sample of grown children still living with both parents. Section 3 reports the main empirical results. The first subsection considers the spousal–BMI relationship in detail, using both a proxy-variable approach and a hierarchical approach to deal with the endogeneity problem caused by omitted variables. The second subsection considers the relationship between parental characteristics (including BMI) and the BMI of grown children (aged 18 and above). Section 4 concludes.

2 Data

This article utilizes data from the 2003 Community Tracking Study (CTS) Household Survey, obtained by the authors through the Inter-University Consortium for Political and Social Research (ICPSR). The CTS (funded by the Robert Wood Johnson Foundation) has existed since 1996, providing longitudinal health-care data from surveys of households and physicians in 60 different communities in the United States. A complete list of the 60 survey "sites" is provided in Appendix A (see Table 5).[3] Within each CTS household surveyed, data were collected on every adult in the household. One adult in the household provided general information for all family members, including insurance coverage, employment, earnings, and race/ethnicity. Answers to detailed health and behavioral questions, such as those regarding smoking, height, and weight, were provided *directly* by each adult in the household. This aspect of the survey minimizes concerns sometimes associated with health and weight measures obtained via proxy responses.[4]

The original CTS data contain information on 46,587 individuals from 25,419 families. (A "family" can consist of a single person.) To focus upon familial relationships, we limit ourselves to the subsample of families containing a married couple. To avoid issues associated with survival bias and potential non-linearities of the BMI profile at later ages, we further restrict our subsample to include only married couples for which both spouses are under the age of 65. The first two columns of Table 1 provide

[3] Although the CTS has a longitudinal aspect to it, household respondents were not asked about their height and weight status until the 2003 survey.

[4] The user's guide published by the Center for Studying Health System Change (2005) provides further details on the sample design and the household survey.

Table 1 Variables and descriptive statistics

	Married sample		Sample of grown children (ages 18–30)	
	Female	Male	Female	Male
Non-indicator variables				
BMI (weight/height^2(kg/m^2))	26.32	27.99	23.84	25.02
	(5.41)	(4.37)	(5.11)	(4.62)
Family income (in $)	70,974	70,974	64,320	58,404
	(38,887)	(38,887)	(49,353)	(48,798)
Education (in years)	13.98	14.05	13.27	12.83
	(2.44)	(2.60)	(1.74)	(1.71)
Age	44.01	46.07	21.06	21.02
	(10.00)	(10.16)	(2.93)	(2.86)
Indicator variables				
Obese (1 if BMI $\geq$ 30)	0.2271	0.2795	0.1205	0.1419
Overweight (1 if BMI $\geq$ 25)	0.5271	0.7620	0.3163	0.4237
Has child	0.5317	0.5317	0.4032	0.2710
No health insurance	0.0798	0.0788	0.1217	0.1871
Smoker	0.1776	0.2146	0.1808	0.2731
Employed	0.6506	0.8224	0.6199	0.6355
White	0.8035	0.8084	0.8030	0.8280
Black	0.0631	0.0685	0.0765	0.0581
Hispanic	0.0892	0.0811	0.0881	0.0613
Other race	0.0442	0.0421	0.0324	0.0527
Married	1.0000	1.0000	0.5979	0.5086
Number of observations	7055	7055	863	930

Standard deviations reported in parentheses for non-indicator variables

descriptive statistics for the wives and husbands from this subsample. A total of 7,055 families with a married couple had complete information for the variables of interest and were retained in the data. The table provides sample averages for the variables, along with standard deviations for the non-indicator variables. The BMI variable (units kg/m^2) is constructed by dividing the respondent's weight by the square of the respondent's height. This data item, but not height and weight individually, is observed.[5]

Two variables in Table 1, "family income" and "has child," are common to the spouses within a household, with the other variables possibly taking on different values for husbands and wives. The "family income" variable is top-coded at $150,000, resulting in 8.4% of the married-sample observations being top-coded. The "smoking"

[5] In the data, the BMI values are bottom-coded at 18 and top-coded at 40, with roughly 0.9% of observations bottom-coded and 2.8% of observations top-coded. Due to the minimal amount of censoring involved, BMI is treated as a continuous variable throughout this article. Application of censored regression techniques yielded no substantive differences in the empirical results.

indicator variable is equal to one for individuals who reported smoking either "everyday" or "some days." The "employed" indicator variable is equal to one for individuals who reported working either for pay or profit during the week prior to the interview. Looking at the first two columns of Table 1, the average BMI of men is larger than that of women, whereas the variation in BMI is greater for women than men. For the married sample, roughly a quarter of men and women are classified as obese (BMI $\geq$ 30). A far greater percentage of men (76.2%) than women (52.7%), however, are classified as being overweight (BMI $\geq$ 25). Roughly 8% of married individuals in the sample have no health insurance. Men are more likely to be smokers than women (21.5% vs. 17.8%) and more likely to be employed than women (82.2% vs. 65.1%). The racial composition of the married sample is roughly 80–81% white, 6–7% black, 8–9% Hispanic, and 4% other races.

The last two columns of Table 1 provide descriptive statistics for the sample of grown children (ages 18–30) that still live in the same household as both their mother and father. This subsample consists of 863 females and 930 males.[6] Although it is possible that grown children living with their parents differ in unobservable ways from those not living with their parents, the data only allow us to observe parental information for the former group. We focus upon younger grown children (ages 18–30) since this selection issue is likely exacerbated at later ages. In Appendix B, we provide evidence that the two samples of grown children (living with parents and living without parents) are quite comparable in terms of their conditional BMI distributions; although we cannot rule out unobservable differences, these results provide greater confidence that our sample is representative of the overall sample of 18–30-year-olds.

Before undertaking regression analyses (Sect. 3), we conclude this section by briefly considering the (unconditional) relationship between husband BMI and wife BMI. The correlation in spousal BMI is equal to 0.266. Although a correlation value of 0.266 may appear low at first glance, BMI is highly variable and spousal BMI is quite clearly one of its strongest predictors. As a comparison, for the wives in the sample, the correlation of BMI with own education is -0.165, with own smoking indicator is -0.039, and with log-family income is -0.150. For the husbands in the sample, the correlation of BMI with own education is -0.096, with own smoking indicator is -0.092, and with log-family income is -0.020.

Figures 3 and 4 provide a graphical view of the spousal–BMI relationship. Figure 3 plots the 25% percentile, median, and 75% percentile of husband BMI versus wife BMI, and Fig. 4 reverses the roles of the two quantities. The positive relationship is evident in both plots, and the higher BMI variance among women is reflected in the larger spread between the curves in Fig. 4. Slope estimates from simple linear regressions (reported below in Table 2, specification (1)) indicate that an increase of one unit in husband's BMI is associated with an expected increase of 0.330 units in wife's BMI, whereas an increase of one unit wife's BMI is associated with an expected increase of 0.215 units in husband's BMI.

[6] Unfortunately, height and weight information data were not collected for children under the age of 18.

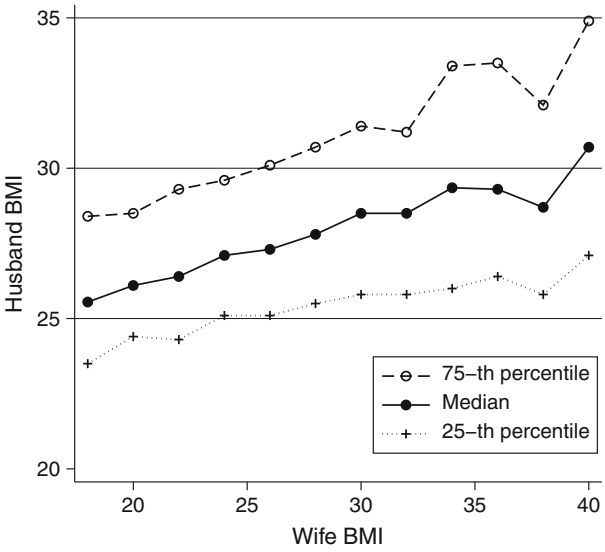

Fig. 3 Spousal–BMI relationship

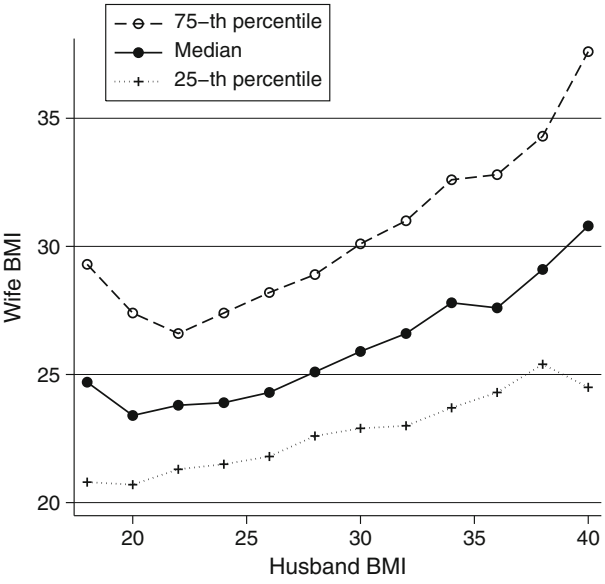

Fig. 4 Spousal–BMI relationship

3 Results

This section is broken into three parts. Section 3.1 examines determinants of individual BMI, using spousal data to control for possible endogeneity. Since omitted variables likely contribute to unobserved heterogeneity in weight outcomes, two alternative

Table 2 BMI regression results for the married sample

	Dep. var. = wife BMI				Dep. var. = husband BMI			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Spouse BMI	0.330**		0.289**	0.285**	0.215**		0.196**	0.195**
	(0.016)		(0.016)	(0.016)	(0.010)		(0.011)	(0.011)
Spouse education				−0.208**				−0.098**
				(0.030)				(0.027)
Spouse employed				−0.176				0.236**
				(0.180)				(0.109)
log(family income)		−1.807**	−1.842**	−1.400**		0.204	0.463**	0.563**
		(0.276)	(0.274)	(0.271)		(0.207)	(0.207)	(0.216)
Has child		−0.417**	−0.369**	−0.339**		−0.373**	−0.292**	−0.264**
		(0.159)	(0.153)	(0.153)		(0.123)	(0.118)	(0.118)
No health		−0.084	−0.017	−0.095		−0.190	−0.163	−0.172
insurance		(0.264)	(0.257)	(0.255)		(0.221)	(0.215)	(0.216)
Smoker		−1.119**	−0.981**			−1.343**	−1.258**	
		(0.174)	(0.168)			(0.131)	(0.128)	
Smoker, spouse				−0.679**				−1.233**
non-smoker				(0.244)				(0.162)
Smoker, spouse				−1.188**				−1.228**
smoker				(0.212)				(0.175)
Non-smoker,				0.687**				0.472**
spouse smoker				(0.216)				(0.198)
Education		−0.276**	−0.220**	−0.110**		−0.201**	−0.143**	−0.093**
		(0.030)	(0.029)	(0.033)		(0.023)	(0.022)	(0.026)
Age		0.152**	0.135**	0.136**		0.183**	0.156**	0.139**
		(0.052)	(0.051)	(0.051)		(0.044)	(0.043)	(0.043)
Age^2		−0.0011*	−0.0011*	−0.0011*		−0.0018**	−0.0016**	−0.0014**
		(0.0006)	(0.0006)	(0.0006)		(0.0005)	(0.0005)	(0.0005)
Employed		0.230*	0.141	0.069		−0.166	−0.138	−0.163
		(0.138)	(0.134)	(0.134)		(0.152)	(0.149)	(0.149)
Black		2.476**	2.323**	2.260**		0.583**	0.118	0.125
		(0.281)	(0.273)	(0.270)		(0.232)	(0.226)	(0.227)
Hispanic		0.531**	0.592**	0.477*		0.441**	0.228	0.193
		(0.253)	(0.246)	(0.245)		(0.217)	(0.212)	(0.214)
Other race		−0.760**	−0.352	−0.297		−0.949**	−0.997**	−0.974**
		(0.299)	(0.298)	(0.297)		(0.257)	(0.239)	(0.239)
Site dummies?	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Table 2 continued

	Dep. var. = Wife BMI				Dep. var. = Husband BMI			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
R squared	0.071	0.086	0.138	0.147	0.071	0.053	0.107	0.110
Number of observations	7055	7055	7055	7055	7055	7055	7055	7055

* Significant at 10% level; ** significant at 5% level. Heteroskedasticity-robust standard errors, clustered at the couple level, are reported in *parentheses*. The omitted racial category is "White." The omitted smoking category in specification (4) is both non-smokers (i.e.,"non-smoker, spouse non-smoker")

approaches of dealing with the endogeneity are considered: a proxy-variable approach (using spousal characteristics as proxy variables) and a hierarchical approach (allowing for a shared "couple" effect). Section 3.2 examines the relationship between grown children's BMI and observable characteristics (including BMI) of their parents.[7]

3.1 Spousal BMI models

Letting the subscript h denote "husband," we consider the following model for husband's BMI:

$$\text{BMI}_\text{h} = x_\text{h}\beta_\text{h} + c_\text{h} + u_\text{h}. \tag{1}$$

The (row) vector x_h consists of observable covariates (with associated coefficient vector β_h). The last two terms on the right-hand-side of Eq. 1 are unobservables. u_h is an idiosyncratic error disturbance, assumed to satisfy $E(u_\text{h}|x_\text{h}) = 0$. On the other hand, c_h may be correlated with x_h and, therefore, may introduce an endogeneity problem. The term c_h arises in the weight-outcome equation due to omitted variables, such as measures of exercise and nutrition, and may cause inconsistency of β_h estimated obtained by ordinary least squares (OLS). As an example, consider the association of BMI with education. If more educated individuals are more likely to exercise and have healthy eating habits, we would expect that the education measure (years of education) would be negatively correlated with the unobservable c_h. Therefore, the OLS estimate of the education coefficient would have a negative omitted-variables bias (i.e., an estimate lower than the causal effect of education on BMI). As discussed below, we will utilize proxy variables and a hierarchical model in order to ameliorate the omitted-variables bias arising from the unobserved heterogeneity c_h.

An analogous model for wife's BMI, with subscript w denoting "wife," is given by

$$\text{BMI}_\text{w} = x_\text{w}\beta_\text{w} + c_\text{w} + u_\text{w}, \tag{2}$$

with the observables denoted x_w and the unobservables denoted c_w and u_w (with $E(u_\text{w}|x_\text{w}) = 0$). Note that the coefficient vector β_w may differ from β_h, allowing

[7] Qualitatively similar results have been obtained using obesity and overweight indicators as dependent variables. These results are available upon request from the authors.

observables to have potentially different marginal effects on BMI for husbands and wives.

To deal with the possible endogeneity in the models in Eqs 1 and 2, two alternative (but related) strategies are considered. The first strategy is to include spousal variables (specifically BMI, smoking status, education, and employment status) as proxies for the omitted variables. For instance, if p_w denotes the vector of proxy variables for a wife, the proxy-variable regression for husband's BMI would regress BMI_h upon x_h and p_w. Similarly, the proxy-variable regression for wife's BMI would regress BMI_w upon x_w (wife's observables) and p_h (husband's proxy variables). This proxy-variable strategy may also be viewed as a predictive-regression exercise in which spousal control variables are used to offer more accurate predictions of BMI. The second strategy is to view the spousal data in a hierarchical framework, where the the existence of data on both spouses allows one to control for an unobserved, common "couple effect" that influences both BMI values.[8]

3.1.1 Proxy-variable regressions

The main results for the proxy-variable regressions are reported in Table 2. The table has two panels, one with wife's BMI as the dependent variable and the other with husband's BMI as the dependent variable. Each panel reports the coefficient estimates from four different regression specifications: (1) simple linear regression of BMI on spousal BMI, (2) linear regression of BMI on individual characteristics (x_w or x_h), (3) linear regression of BMI on individual characteristics and spousal BMI, and (4) linear regression of BMI on individual characteristics and a larger set of spousal proxy variables (BMI, smoking status, education, and employment status). For specifications (2), (3), and (4), each of the individual explanatory variables from Table 1 is included in the regressions. The family income variable is included in logarithmic form, and a quadratic specification for age is used. The excluded racial category is "white," so the coefficient estimates on the other racial indicator variables should be interpreted as differences from whites. Finally, a set of site dummies (to control for any possible differences across the 60 CTS geographical locations) is also included for these specifications. Site-dummy estimates for specification (4) are provided in Appendix A (Table 5).

Spousal BMI as a predictor of BMI: A consistent finding in Table 2 is the importance of spousal BMI as a predictor of BMI for both men and women. Across all specifications, the t-statistic for spousal BMI is larger than any of the t statistics for the individual variables or the other spousal variables. By itself (specification (1)), spousal BMI explains 7.1% of the variation in BMI. As a comparison, the individual characteristics of specification (2) (without site dummies) explain 7.1% of the variation in women's BMI and only 3.5% of the variation in men's BMI. Although we retain them for robustness, the site dummies explain very little of the variance in

[8] We call this a "hierarchical model," although other terminology ("matched-pair model," "cluster model," etc.) is commonly used for this situation (see, for example, Wooldridge 2002, Section 11.5).

BMI outcomes.[9] Importantly, spousal BMI remains a strong predictor of an individual's BMI even after controlling for observable characteristics. For explaining wife's BMI, the coefficient estimate on spousal BMI decreases from 0.330 to 0.286 (a 13.3% change) when comparing specifications (1) and (4); for explaining husband's BMI, the decrease is from 0.215 to 0.196 (a 8.8% change).[10] Even with the significant effects found for the other included spousal variables in specification (4), these additional variables have essentially no effect on the estimated association with spousal BMI (comparing specifications (3) and (4)).

Association of BMI with individual characteristics: Without spousal controls (specification (2)), the presence of children, smoking, and education has significant negative associations with BMI for both men and women. The magnitude of the coefficient estimates for these three variables is also very similar for men and women. Health insurance has no significant relationship with BMI for either men or women. The age profile has a similar shape for men and women, with the estimates implying a declining positive association of age with BMI as age increases.[11] Blacks and Hispanics, on average, have higher BMI values than whites after controlling for other observables; a more pronounced difference (2.48 units of BMI) is estimated among black women as compared to black men (0.58 units of BMI). A major difference between the results for wives and husbands concerns the association of BMI with family income. There is a very significant negative association for wives, whereas the association is insignificant for husbands. For wives, the coefficient estimate implies that an increase of 1% in family income is associated with an expected decrease of 0.018 in BMI. This finding is consistent with previous study that has found a negative relationship between wages and obesity among women but not men (Cawley 2004; Conley and Glauber 2006).[12]

Association of BMI with other spousal variables: The inclusion of spousal proxy variables in specification (4) yields several interesting results. Even after controlling for spousal BMI, spousal smoking status and spousal education are statistically significant predictors of an individual's BMI. For women in the sample, expected BMI is -0.208 units (t stat $= 6.9$) lower for each additional year of husband's education; for men in the sample, the association is somewhat lower in magnitude (-0.098 units for a year of wife's education) but still very statistically significant (t stat $= 3.6$). A negative association between BMI and spousal education would be expected if spousal

[9] The 60 site dummies, by themselves, explain only 2.1% of the variance in women's BMI and 1.8% of the variance in men's BMI. In specifications (2)–(4), their inclusion has little effect on either overall R-squared or the coefficient estimates on the other included variables. In fact, a partial F test on the site dummies for these specifications would not reject the null of no site heterogeneity.

[10] Another way to see the persistence in the spousal–BMI relationship is to consider the estimated residuals from specification 2. Using the estimated residuals from these regressions, the correlation between estimated residuals for wives and husbands is equal to 0.231 (as compared to the original BMI correlation value of 0.266).

[11] The marginal effect of age turns negative at around 69 years of age for women and 51 years of age for men.

[12] Cawley (2004), using NLSY data, argues that higher weight has a causal impact by lowering wages for white females. Conley and Glauber (2006), using sibling data from the Panel Study of Income Dynamics (PSID), find a negative effect of a BMI on family income for women but no such effect for men.

education is negatively correlated with c_h (or c_w). Specifically, higher spousal education may be proxying for unobservable characteristics like better exercise and healthy eating.[13] To investigate the impact of spousal smoking upon BMI in specification (4), we constructed four categories based upon the interaction of own smoking status and spousal smoking status; the baseline (omitted) category is both being non-smokers.[14] For non-smokers, there is a positive association of BMI with spousal smoking status. Women (men) have an expected BMI 0.687 (0.472) units higher if their spouse smokes. This finding is consistent with spousal smoking proxying for one's own poor exercise/nutritional habits. Interestingly, however, this association does not appear among smokers. Looking at the results for men, smokers have an expected BMI about 1.23 units lower than the baseline category, whether or not their spouse smokes; the expected-BMI difference from the non-smoker/spouse-smoker category is larger (1.70 units lower). For women, the lowest expected BMI (holding all else fixed) is found in the both-smokers category (1.19 units lower than the baseline, 1.88 units lower than the non-smoker/spouse-smoker category). Overall, these results support the notion that (own) smoking has a negative association with BMI due to physiological reasons.[15]

Impact of proxy variables upon coefficient estimates: Comparing specifications (2) and (4) to each other, the inclusion of spousal variables leads to a decrease in magnitude for several of the coefficient estimates, including the variables for presence of children and education and the indicators for blacks and Hispanics. The largest change occurs for the education variable, where the magnitude of the coefficient estimate drops by more than 50% for both women (from -0.276 to -0.111) and men (from -0.201 to -0.093). This large change suggests that the original education-coefficient estimate from specification (2) was subject to an important omitted-variables bias. The asymmetric association (depending upon gender) of BMI with income remains after inclusion of spousal variables. The relationship remains significantly negative for women, with the coefficient estimate on log(family income) changing from -1.807 to -1.401. For men, the association changes from a statistically insignificant one to a significantly *positive* relationship, with a coefficient estimate of 0.563 on log(family income).[16]

[13] We were unable to find previous research that actually documents a correlation between own exercise/nutrition and spousal education. On the other hand, the correlation between own exercise/nutrition and own education has been well documented, and we have no reason to suspect a similar (but smaller in magnitude) correlation would not exist with spousal education.

[14] We are grateful to a referee for suggesting this interaction approach.

[15] It is widely believed that smoking suppresses appetite, but estimation of this "causal" effect has resulted in mixed findings. Chou et al. (2004) find a positive elasticity of BMI (and obesity) with respect to cigarette prices, suggesting a negative relationship between smoking and body weight. Cawley et al. (2004) find that smoking initiation is more likely among female teenagers who are overweight or trying to lose weight. In contrast, Gruber and Frakes (2006) find a negative elasticity of BMI (and obesity) with respect to cigarette taxes, suggesting a *positive* relationship between smoking and weight.

[16] We examined whether the BMI-income association varied by employment status and found no significant differences between labor-force participants and non-participants. Thus, there is no strong evidence from this data that labor-force concerns (e.g., wage penalty for obesity) or time-use issues (e.g., less time for food preparation) are driving the association.

3.1.2 Hierarchical-model analysis

To estimate how the unobservable common (couple) component in spousal BMI values are related to observables (x_h and x_w), we use the "correlated random-effects" estimator introduced by Chamberlain (1982).[17] First, we explicitly assume that the "couple effect" (denoted c) is common to a husband and wife. In terms of the BMI models in (1) and (2) above, we have $c = c_h = c_w$, so that

$$\mathrm{BMI}_h = x_h\beta_h + c + u_h \tag{3}$$

$$\mathrm{BMI}_w = x_w\beta_w + c + u_w. \tag{4}$$

Note that the model remains flexible with respect to the marginal effects of the observables upon the BMI values since β_h and β_w are not assumed to be the same.[18]

The *correlated random-effects model* of Chamberlain (1982) views the unobservable c as a linear projection onto the observables x_h and x_w:

$$c = \psi + x_h\lambda_h + x_w\lambda_w + v, \tag{5}$$

where ψ is a scalar and v is a disturbance that (by definition of linear projections) is uncorrelated with x_h and x_w. Combining Eqs. 3 and 4 with the linear projection in (5) yields

$$\mathrm{BMI}_h = \psi + x_h\beta_h + x_h\lambda_h + x_w\lambda_w + (v + u_h) \tag{6}$$

$$\mathrm{BMI}_w = \psi + x_w\beta_w + x_h\lambda_h + x_w\lambda_w + (v + u_w) \tag{7}$$

The simplest way to estimate the parameters (ψ, β_h, β_w, λ_h, λ_w) is to use pooled least-squares regression (see, e.g., Wooldridge 2002, Section 11.3), which is the approach taken here. An important benefit of the CRE estimator is that, unlike the fixed-effects estimator, the analysis (specifically, the λ_h and λ_w estimates) directly reveals which of the observable variables are correlated with the unobservable component c.[19] For elements of x_h and x_w that do not vary for wives and husbands (specifically, log(familyincome) and the Has child indicator), note that the pooled regression can not separately identify the associated components of β_h and λ_h (or the associated components of β_w and λ_w).

The results from CRE estimation are reported in Table 3. For comparison, the first two columns replicate the specification (2) estimates of Table 2. The last four columns report the CRE estimation results, with β_w, β_h, λ_w, and λ_h reported from

[17] A "pure" random-effects model, in which the common effect is assumed to be independent of the observables, is overwhelmingly rejected by the standard Hausman test.

[18] While we also estimated a version of this model with a differential effect of c upon husbands and wives, the results were extremely similar and are not reported here.

[19] The pooled least-squares regression yields numerically identical estimates of β_h and β_w to the more commonly used fixed-effects ("within") regression. Equivalently, one could estimate use the fixed-effects estimator to estimate the fixed effect for each couple and then project these estimated fixed effects upon x_h and x_w to estimate λ_h and λ_w.

Table 3 Hierarchical-model BMI regression results

	OLS regressions		Correlated random-effects regression			
			β Estimates		λ Estimates	
	Wife	Husband	Wife	Husband	Wife	Husband
Male				−6.319**		
				(1.737)		
log(family income)	−1.807**	0.204	−1.283**	0.278		
	(0.276)	(0.207)	(0.269)	(0.214)		
Has child	−0.417**	−0.373**	−0.384**	−0.214*		
	(0.159)	(0.123)	(0.159)	(0.128)		
No health insurance	−0.084	−0.190	−0.194	−0.148	−0.222	0.168
	(0.264)	(0.221)	(0.469)	(0.468)	(0.324)	(0.395)
Smoker	−1.119**	−1.343**	−1.218**	−1.314**	0.026	−0.068
	(0.174)	(0.131)	(0.215)	(0.203)	(0.152)	(0.180)
Education	−0.276**	−0.201**	−0.020	0.110**	−0.129**	−0.250**
	(0.030)	(0.023)	(0.038)	(0.035)	(0.027)	(0.031)
Age	0.152**	0.183**	0.250**	0.218**	−0.111	0.018
	(0.052)	(0.044)	(0.098)	(0.097)	(0.071)	(0.082)
Age^2	−0.0011*	−0.0018**	−0.0029**	−0.0028**	0.0017**	0.0001
	(0.0006)	(0.0005)	(0.0011)	(0.0010)	(0.0008)	(0.0009)
Employed	0.230*	−0.166	−0.167	0.010	0.307**	−0.197
	(0.138)	(0.152)	(0.156)	(0.215)	(0.112)	(0.184)
Black	2.476**	0.583**	1.588**	−0.304	−0.051	0.944
	(0.281)	(0.232)	(0.687)	(0.673)	(0.478)	(0.600)
Hispanic	0.531**	0.441**	0.490	0.085	−0.697**	0.821**
	(0.253)	(0.217)	(0.327)	(0.365)	(0.247)	(0.331)
Other race	−0.760**	−0.949**	0.148	−0.914**	−1.221**	0.714*
	(0.299)	(0.257)	(0.421)	(0.390)	(0.295)	(0.390)

* Significant at 10% level; ** significant at 5% level. Heteroskedasticity-robust standard errors are reported in *parentheses*. The omitted racial category is "White." Site dummies are included in all regressions

left to right. The log(family income) and Has child estimates are reported in the β_w and β_h columns, but these estimates should be interpreted as the estimated *overall* effects of these two variables as measured by the associated components of $\hat{\beta}_w + \hat{\lambda}_w + \hat{\lambda}_h$. The *R* squared value for the CRE regression is equal to 10.5%, meaning that 10.5% of the total variation in a married individual's BMI is explained by the hierarchical model. As a comparison, recall that 8.6 and 5.3% of the variations in female and male BMI, respectively, were explained by OLS regressions that did not account for the hierarchical couple effect. For the effects of the individual variables in the CRE regression, we focus the discussion on what we consider the three most interesting: smoking status, education, and income. We discuss each of the three in turn:

Table 4 BMI regression results for grown children (ages 18–30)

	Dep. var. = daughter BMI				Dep. var. = son BMI			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Mother: BMI	0.304**		0.258**	0.253**	0.222**		0.211**	0.224**
	(0.032)		(0.031)	(0.032)	(0.030)		(0.029)	(0.029)
Mother: smoker				0.714				0.836**
				(0.532)				(0.421)
Mother: education				0.064				−0.107
				(0.076)				(0.075)
Mother: employed				−0.638*				0.763**
				(0.365)				(0.329)
Father: BMI	0.259**		0.255**	0.267**	0.180**		0.170**	0.173**
	(0.045)		(0.044)	(0.044)	(0.037)		(0.037)	(0.037)
Father: smoker				1.061**				0.867**
				(0.443)				(0.387)
Father: education				−0.053				−0.099
				(0.072)				(0.070)
Father: employed				−0.806				0.100
				(0.492)				(0.403)
log(family income)		−0.630	−0.271	−0.123		0.255	0.432	0.656*
		(0.424)	(0.401)	(0.403)		(0.347)	(0.340)	(0.340)
Has child		0.226	0.233	0.170		0.029	0.231	0.328
		(0.388)	(0.355)	(0.357)		(0.411)	(0.382)	(0.381)
No health insurance		0.176	−0.197	−0.310		−0.059	−0.194	−0.235
		(0.581)	(0.530)	(0.530)		(0.439)	(0.424)	(0.419)
Smoker		0.732	0.825*	0.514		−0.782**	−0.662**	−0.756**
		(0.476)	(0.442)	(0.436)		(0.332)	(0.324)	(0.327)
Education		−0.314**	−0.184	−0.138		−0.136	−0.052	0.072
		(0.133)	(0.127)	(0.131)		(0.110)	(0.102)	(0.113)
Age		−1.272	−1.087	−1.200		0.579	0.230	0.139
		(0.895)	(0.838)	(0.830)		(0.778)	(0.762)	(0.774)
Age^2		0.037*	0.030	0.032*		−0.005	0.001	0.003
		(0.019)	(0.018)	(0.018)		(0.017)	(0.017)	(0.017)
Employed		−0.266	−0.392	−0.386		−0.579*	−0.646**	−0.782**
		(0.351)	(0.322)	(0.325)		(0.331)	(0.312)	(0.312)
Black		2.029**	1.698**	1.530**		0.798	0.383	0.602
		(0.674)	(0.636)	(0.632)		(0.764)	(0.742)	(0.717)
Hispanic		0.683	0.428	0.435		1.374*	1.237*	0.993
		(0.632)	(0.591)	(0.616)		(0.764)	(0.664)	(0.655)
Other race		1.141	0.816	0.817		−1.546**	−1.298**	−1.263**
		(1.143)	(1.023)	(0.955)		(0.643)	(0.630)	(0.638)
Married		−0.161	−0.338	−0.289		−0.349	−0.544	−0.395
		(0.614)	(0.572)	(0.577)		(0.539)	(0.523)	(0.519)

Table 4 continued

	Dep. var. = Daughter BMI				Dep. var. = Son BMI			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
R squared	0.188	0.094	0.235	0.251	0.119	0.058	0.162	0.191
Number of observations	863	863	863	863	930	930	930	930

* Significant at 10% level; ** significant at 5% level. Heteroskedasticity-robust standard errors are reported in *parentheses*. The omitted racial category is "White."

The association between smoking and BMI: The estimated effect of smoking on expected BMI in the hierarchical model is −1.218 for women and −1.314 for men. Both effects are statistically significant with t statistics around 6. The magnitude of the estimates are quite close to those found in both the OLS regressions without spousal control variables and also the spousal control regressions. The relationship of the couple effect with smoking status (as measured by the λ estimates) is found to be statistically insignificant. Overall, these findings (i) support the physiological argument that smoking has a negative association with BMI and (ii) indicate that smoking does not appear to have an indirect effect upon BMI that works through the shared couple effect.

The association between education and BMI: The CRE estimates of the education coefficients indicate that the original OLS estimates were severely downward biased due to omitted variables. Both wife's education and husband's education have a significant negative association with c, in line with the negative association found between BMI and spousal education in Table 2. Interestingly, once the couple effect is controlled for, the causal effect of education on BMI for women (−0.020 with a standard error of 0.038) is statistically insignificant. This estimate suggests that the causal impact of education upon BMI is minimal, although education itself is strongly correlated with the couple-effect component of BMI. Strangely, the estimate of education's effect on BMI is positive for men. While we don't have a great explanation for why such a relationship might exist, we note several issues. First, the overall association between education and BMI for men is still negative (combining the 0.110 and −0.250 estimates). Second, this positive association may not be too surprising given the positive association between income and BMI (for males) seen in the other estimation results.

The association between income and BMI: Although the β and λ estimates are not separately identified for family income, the CRE estimates do confirm a significant *difference* between the effects of family income on wife's BMI and husband's BMI. The overall effect on wife's BMI is significantly negative whereas the overall effect on husband's BMI is positive but statistically insignificant. The overall estimated effect of log(family income) (−1.283) for wives is fairly close to the proxy-variable estimate (−1.400) found in the proxy-variable regression (Table 2, specification (4)).

3.2 BMI regressions for grown children

In this section, we analyze the intergenerational transmission of BMI based upon the CTS data on grown children (ages 18–30) still living with both of their

parents.[20] The sample of grown children consists of 863 women and 930 men, with descriptive statistics provided in Table 1 of Sect. 2. As compared to the sample of married individuals analyzed above, this sample is younger (with less prevalence of obesity and overweight), less educated, and more likely to lack health insurance; among men, they are also significantly less likely to be employed and more likely to smoke.

The proxy-variable regression approach of Sect. 3.1.1 is used to model both daughter's BMI and son's BMI. Rather than proxying omitted variables by a spouse's characteristics, however, the omitted variables are proxied by the characteristics of both parents. Table 4 reports the regression results. Specifications (1) are (2) are the baseline specifications, with (1) including parents' BMI as covariates and (2) including individual characteristics as covariates. Specifications (3) and (4) are the proxy-variable regressions, with (3) using only parents' BMI values as proxies and (4) using the the larger set of parental proxy variables (BMI, smoking status, education, employment status).

Parental BMI as a predictor of BMI: Parental BMI values are extremely important predictors of an individual's BMI. From specification (1), note that parental BMI alone explains 18.8 and 11.9% of female BMI and male BMI, respectively. These R-squared values are roughly *twice* as large as those found in specification (2), where only individual characteristics (but not parental BMI) are used in the regression. Looking at the most complete model in specification (4), a one-unit increase in mother's BMI (father's BMI) is associated with an expected increase in daughter's BMI of 0.253 units (0.267 units), and a one-unit increase in mother's BMI (father's BMI) is associated with an expected increase in son's BMI of 0.224 units (0.173 units). Similar to our finding in the spousal data, inclusion of individual characteristics in the regression specifications of Table 4 has only a small effect on the estimated coefficients on mother's BMI and father's BMI.

Parental smoking and BMI: Given the comparatively small sample size here (roughly ten times smaller than the spousal sample considered previously), the standard errors on the own-smoking coefficients are quite large. We do, however, find a significant negative association between smoking and BMI among males (-0.756 units in specification (4)). For females, the own-smoking coefficient in specification (4) is statistically insignificant. For both sons and daughters, the coefficient estimates for the parental smoking indicator variables are all significantly positive, with magnitudes ranging between 0.714 and 1.061; three of the four estimates are significant at a 5% level, with only the mother-smoking variable being statistically insignificant in the daughter BMI regression. As in our interpretation of the spousal smoking variable for the married sample, we believe that parental smoking is proxying for behavior and attitudes that are positively correlated with BMI (such as poor diet or lack of exercise).

Parental income and BMI: The "family income" variable is a measure of parental income for this sample of grown children. Therefore, the estimates in Table 4 quantify the association between BMI and parental income. The estimates here are a bit imprecise due to the small sample size, but the signs of the estimates are in agreement with

[20] As discussed in Sect. 2, to check that this sample is representative of the overall young-adult population, the appendix provides a comparison of this sample to a sample of 18–30-year-olds not living with their parents.

the asymmetric findings in the spousal models. Across all specifications, the income coefficient for females is negative (but statistically insignificant) whereas the income coefficient for males is positive (significant at a 10% level specification (4)).

4 Conclusion

This study has utilized recent household survey data in order to study familial relationships in weight outcomes, controlling for socioeconomic, behavioral, and demographic characteristics of both the individual and the individual's family members. Spousal BMI (in the married sample) and parental BMI (in the grown-child sample) are found to be the most important predictors of an individual's BMI, even after controlling for individual characteristics such as education, income, smoking, etc. In a hierarchical model that allowed for a shared couple effect, the estimates suggest that smoking has a negative causal relationship with BMI. In contrast, the negative association of BMI with education disappears in the hierarchical model, suggesting that education does not have a negative causal impact upon BMI (but rather proxies for other unobservables).

While this study sheds some light on familial correlations in BMI and begins to tackle some of the important endogeneity issues, we are ultimately limited by the cross-sectional nature of the data. For instance, to better understand the causal mechanisms underlying the positive relationship between spousal/parental smoking and individual BMI, it would be ideal to have a longitudinal dataset that tracks changes in smoking status and BMI (along with other demographic characteristics) for all family members. Such data, especially if augmented with detailed information on exercise habits and nutritional choices, will offer economists the opportunity to clarify causal channels by which BMI is affected and, therefore, better inform public policy aimed at reducing obesity.

Acknowledgments The Community Tracking Study (CTS) data were obtained from the Inter-University Consortium for Political and Social Research (ICPSR). The authors are grateful to seminar participants at IUPUI and the 13th Annual Panel Data Conference for their comments.

Appendix A: CTS survey sites

The CTS household data for 2003 was collected in 60 different communities across the United States. The sites are listed in Table 5. According to the CTS user's guide published by the Center for Studying Health System Change (2005), "[t]he sites generally conform to the metropolitan statistical areas (MSAs) defined by the Office of Management and Budget and the non-metropolitan portions of the economic areas defined by the Bureau of Economic Analysis." In the table, therefore, a site labeled by a city's name may be associated with more than one state; for example, the Augusta site has households in both Georgia and South Carolina. Metcalf et al. (1996) provides additional details on the definitions of the CTS sites. To give a sense of the sample composition, Table 5 reports the number of households for each site within the full married sample of 7,055 households. In addition, we provide the de-meaned site-dummy coefficient estimates for the wife-BMI and husband-BMI regressions reported in Table 2.

Table 5 Community Tracking Study (CTS) Sites

Geographic area	Number of observations in married sample	Percent	Demeaned coefficient wife BMI	Demeaned coefficient husband BMI
Huntington (WV/KY/OH)	98	1.39	1.38	0.24
Middlesex (NJ)	88	1.25	1.09	−0.10
Detroit (MI)	87	1.23	0.99	0.14
N Georgia	79	1.12	0.98	0.11
Atlanta (GA)	86	1.22	0.94	−1.05
Pittsburgh (PA)	74	1.05	0.87	0.16
Nassau (NY)	81	1.15	0.81	0.08
Portland (OR/WA)	95	1.35	0.73	0.27
Lansing (MI)	269	3.81	0.71	0.20
Killeen (TX)	74	1.05	0.69	0.18
Worcester (MA)	92	1.30	0.69	0.96
E North Carolina	85	1.20	0.64	−0.29
Augusta (GA/SC)	77	1.09	0.60	−1.05
Columbus (OH)	116	1.64	0.54	−0.03
Dothan (AL)	87	1.23	0.48	0.07
Terre Haute (IN)	81	1.15	0.43	0.87
Syracuse (NY)	307	4.35	0.43	0.51
St. Louis (MO/IL)	105	1.49	0.42	0.29
NE Indiana	101	1.43	0.42	0.64
NW Washington	70	0.99	0.39	−0.62
Cleveland (OH)	263	3.73	0.33	−0.42
Denver (CO)	110	1.56	0.33	−0.65
Philadelphia (PA/NJ)	75	1.06	0.30	−0.05
W-Cen Alabama	89	1.26	0.29	0.75
Seattle (WA)	223	3.16	0.29	−0.48
Baltimore (MD)	84	1.19	0.27	0.99
Modesto (CA)	82	1.16	0.25	0.58
Santa Rosa (CA)	81	1.15	0.18	−1.82
Cen Arkansas	112	1.59	0.16	0.73
Riverside (CA)	62	0.88	0.10	−0.41
Indianapolis (IN)	281	3.98	0.05	0.05
Minneapolis (MN/WI)	111	1.57	0.01	−0.53
Rochester (NY)	124	1.76	−0.09	0.32
Milwaukee (WI)	84	1.19	−0.10	0.35
Shreveport (LA)	68	0.96	−0.10	0.66
Little Rock (AR)	277	3.93	−0.11	0.06
NE Illinois	97	1.37	−0.20	0.44

Table 5 continued

Geographic area	Number of observations in married sample	Percent	Demeaned coefficient wife BMI	Demeaned coefficient husband BMI
Chicago (IL)	76	1.08	−0.22	0.24
Tampa (FL)	70	0.99	−0.25	−0.74
Phoenix (AZ)	216	3.06	−0.27	0.04
Greenville (SC)	273	3.87	−0.31	0.29
Greensboro (NC)	90	1.28	−0.33	0.28
E Maine	100	1.42	−0.33	−0.10
Las Vegas (NV/AZ)	64	0.91	−0.37	−0.46
San Francisco (CA)	46	0.65	−0.50	−1.40
San Antonio (TX)	87	1.23	−0.54	0.48
New York City (NY)	58	0.82	−0.57	−0.45
Wilmington (NC)	77	1.09	−0.57	0.38
N Utah	138	1.96	−0.62	−0.64
Washington (DC/MD)	80	1.13	−0.69	0.13
Houston (TX)	75	1.06	−0.69	0.87
Newark (NJ)	242	3.43	−0.77	−0.38
Miami (FL)	191	2.71	−0.77	−0.18
Tulsa (OK)	97	1.37	−0.83	0.56
Boston (MA)	228	3.23	−0.96	0.01
Bridgeport (CT)	73	1.03	−0.97	−0.11
Orange County (CA)	227	3.22	−0.97	−0.54
W Palm Beach (FL)	38	0.54	−1.08	−0.73
Los Angeles (CA)	50	0.71	−1.08	−1.28
Knoxville (TN)	84	1.19	−1.18	0.28
Total	7,055			

The sites are sorted (descending order) in Table 5 based upon the site-dummy estimate from the wife-BMI regression.

Appendix B: more details on the grown-children sample

This appendix provides a comparison of the adult-sample considered in Sect. 3 (adults aged 18–30 still living with both parents) and the sample of adults not living with their parents. A comparison of observable characteristics is provided in Table 6. Though some significant differences exist between the two samples, concerns about selection bias are only warranted if these differences also affect results concerning the *conditional* BMI distribution. Since we do not observe parental information for the latter sample, we provide a comparison of the BMI regression results using only individual characteristics in Table 7. Despite the differences seen in Table 6, the associations of observables with (conditional) BMI in Table 7 are quite similar.

Table 6 A comparison of 18–30-year-olds living with and without parents

	Women aged 18–30			Men aged 18–30		
	With parents	Without parents	Difference	With parents	Without parents	Difference
	Mean (s.d.) (1)	Mean (s.d.) (2)	Difference (robust s.e.) (1) − (2)	Mean (s.d.) (3)	Mean (s.d.) (4)	Difference (robust s.e.) (3) − (4)
BMI	23.842	25.627	−1.786**	25.016	26.555	−1.538**
	(5.107)	(5.503)	(0.216)	(4.625)	(4.613)	(0.197)
Has child	0.403	0.534	−0.131**	0.271	0.353	−0.082**
	(0.491)	(0.499)	(0.020)	(0.445)	(0.478)	(0.020)
No health insurance	0.122	0.209	−0.088**	0.187	0.302	−0.115**
	(0.327)	(0.407)	(0.015)	(0.390)	(0.459)	(0.018)
Smoker	0.181	0.256	−0.075**	0.273	0.319	−0.046*
	(0.385)	(0.437)	(0.017)	(0.446)	(0.466)	(0.019)
Education	13.268	13.418	−0.150*	12.826	13.159	−0.333**
	(1.735)	(2.379)	(0.081)	(1.707)	(2.541)	(0.089)
Age	21.061	25.708	−4.647**	21.020	25.577	−4.556**
	(2.927)	(3.261)	(0.125)	(2.855)	(3.322)	(0.130)
Employed	0.620	0.623	−0.003	0.635	0.814	−0.179**
	(0.486)	(0.485)	(0.020)	(0.482)	(0.389)	(0.019)
Black	0.076	0.157	−0.081**	0.058	0.134	−0.076**
	(0.266)	(0.364)	(0.012)	(0.234)	(0.341)	(0.012)
Hispanic	0.088	0.157	−0.069**	0.061	0.193	−0.131**
	(0.284)	(0.364)	(0.013)	(0.240)	(0.395)	(0.013)
Other race	0.032	0.074	−0.042**	0.053	0.069	-0.016
	(0.177)	(0.262)	(0.009)	(0.224)	(0.253)	(0.010)
Married	0.598	0.534	0.064**	0.509	0.479	0.029
	(0.491)	(0.499)	(0.020)	(0.500)	(0.500)	(0.021)
Number of observations	863	1862		930	1339	

Table 7 BMI regression results for adults living with and without parents

	Women aged 18–30		Men aged 18–30	
	With parents	Without parents	With parents	Without parents
log(family income)	−0.630	−0.953**	0.255	0.396
	(0.424)	(0.368)	(0.347)	(0.366)
Has child	0.226	0.485*	0.029	0.484
	(0.388)	(0.288)	(0.411)	(0.341)
No health insurance	0.176	−0.697**	−0.059	−0.420
	(0.581)	(0.321)	(0.439)	(0.312)

Table 7 Continued

	Women aged 18–30		Men aged 18–30	
	With parents	Without parents	With parents	Without parents
Smoker	0.732	−0.013	−0.782**	−0.707**
	(0.476)	(0.311)	(0.332)	(0.279)
Education	−0.314**	−0.311**	−0.136	−0.160**
	(0.133)	(0.061)	(0.110)	(0.057)
Age	−1.272	1.202**	0.579	0.559
	(0.895)	(0.562)	(0.778)	(0.542)
Age^2	0.037*	−0.020*	−0.005	−0.008
	(0.019)	(0.011)	(0.017)	(0.011)
Employed	−0.266	0.433	−0.579*	0.070
	(0.351)	(0.268)	(0.331)	(0.335)
Black	2.029**	2.514**	0.798	1.657**
	(0.674)	(0.392)	(0.764)	(0.415)
Hispanic	0.683	0.201	1.374*	0.567*
	(0.632)	(0.363)	(0.764)	(0.333)
Other race	1.141	0.151	−1.546**	−0.523
	(1.143)	(0.477)	(0.643)	(0.513)
Married	−0.161	0.491	−0.349	0.047
	(0.614)	(0.315)	(0.539)	(0.332)
R-squared	0.094	0.066	0.058	0.049
Number of observations	863	1862	930	1339

* Significant at 10% level; ** significant at 5% level

Heteroskedasticity-robust standard errors are reported in *parentheses*. The omitted racial category is "White"

References

Anderson PM, Butcher KF, Levine PB (2003) Maternal employment and overweight children. J Health Econ 22:477–504

Cawley J (2004) The impact of obesity on wages. J Hum Resour 39:452–474

Cawley J, Markowitz S, Tauras J (2004) Lighting up and slimming down: the effects of body weight and cigarette prices on adolescent smoking initiation. J Health Econ 23:293–311

Center for Studying Health System Change (2005) Community Tracking Study 2003 Household Survey Public Use File: user's guide. Technical Publication No. 58. Washington, DC

Chamberlain G (1982) Multivariate regression models for panel data. J Econ 18:5–46

Chou S-Y, Grossman M, Saffer H (2004) An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. J Health Econ 23:565–587

Classen T, Hokayem C (2005) Childhood influences on youth obesity. Econ Hum Biol 3:165–187

Conley D, Glauber R (2006) Gender, body mass and socioeconomic status: new evidence from the PSID. Adv Health Econ Health Serv Res 17:255–280

Cutler DM, Glaeser EL, Shapiro JM (2003) Why have Americans become more obese? J Econ Perspect 17:93–118

Finkelstein EA, Fiebelkorn IC, Wang G (2004) State-level estimates of annual medical expenditures attributable to obesity. Obes Res 12:18–24

Flegal KM, Carroll MD, Ogden CL, Johnson CL (2002) Prevalence and trends in obesity among US adults, 1999–2000. J Am Med Assoc 288:1723–1727

Gruber J, Frakes M (2006) Does falling smoking lead to rising obesity? J Health Econ 25:183–197

Jeffery RW, Rick AM (2002) Cross-sectional and longitudinal associations between body mass index and marriage-related factors. Obes Res 10:809–815

Lakdawalla D, Philipson T (2009) The growth of obesity and technological change. Econ Hum Biol 7: 283–293

Maes HHM, Neale MC, Eaves LJ (1997) Genetic and environmental factors in relative body weight and human adiposity. Behav Genet 27:325–351

Manson JE, Skerrett PJ, Willett WC (2004) Obesity as a risk factor for major health outcomes. In: Bray GA, Bouchard C (eds) Handbook of obesity: etiology and pathophysiology, 2nd edn. pp 813–824. Available at http://books.google.com/books?id=t-Ccv0QGmeAC

Metcalf CE, Kemper P, Kohn LT, Pickreign JD (1996) Site definition and sample design for the Community Tracking Study. Center for Studying Health System Change, Technical Publication No. 1. Washington, DC. Available at http://www.hschange.com/CONTENT/157/157.pdf

Philipson TJ, Posner RA (2003) The long-run growth in obesity as a function of technological change. Perspect Biol Med 46:S87–S107

Price RA, Reed DR, Guido NJ (2000) Resemblance for body mass index in families of obese African American and European American women. Obes Res 8:360–366

Rice T, Despres JP, Daw EW, Gagnon J, Borecki IB, Perusse L, Leon AS, Skinner JS, Wilmore JH, Rao DC, Bouchard C (1997) Familial resemblance for abdominal visceral fat: the HERITAGE family study. Int J Obes 21:1024–1031

Sacerdote B (2004) What happens when we randomly assign children to families? NBER working paper 10894

Stunkard AJ, Sorensen TI, Hanis C, Teasdale TW, Chakraborty R, Schull WJ, Schulsinger F (1986) An adoption study of human obesity. N Engl J Med 314:193–198

Taveras EM, Rifas-Shiman SL, Berkey CS, Rockett HRH, Field AE, Lindsay Frazier A, Colditz GA, Gillman MW (2005) Family dinner and adolescent overweight. Obes Res 13:900–906

Vogler GP, Sorensen TI, Stunkard AJ, Srinivasan MR, Rao DC (1995) Influences of genes and shared family environment on adult body mass index assessed in an adoption study by a comprehensive path model. Int J Obes 19:40–45

Wilson SE (2002) The health capital of families: an investigation of the inter-spousal correlation in health status. Soc Sci Med 55:1157–1172

Wilson FA (2006) Explaining the growth of child obesity in the U. S. University of Chicago Department of Economics, mimeo. Available at http://chess.uchicago.edu/ccehpe/hew_papers/winter_2006/Wilson%202006.pdf

Wolf AM, Colditz GA (1998) Current estimates of the economic cost of obesity in the United States. Obes Res 6:97–106

Wooldridge JM (2002) Econometric analysis of cross section and panel data. MIT Press, Cambridge, MA

World Health Organization (2003) Obesity and overweight: fact sheet. http://www.who.int/dietphysicalactivity/media/en/gsfs_obesity.pdf. Accessed 9 Jan 2005