

# **BMI Trends, Socioeconomic Status, and the Choice of Dataset**

Michael Grabner  
HealthCore, Inc., Wilmington, DE, USA

## **Online Supplementary Material**

## Appendix on Methods and Procedures

### *Data Sources*

The NHANES is available at [www.cdc.gov/nchs/nhanes.htm](http://www.cdc.gov/nchs/nhanes.htm) and the BRFSS at [www.cdc.gov/brfss](http://www.cdc.gov/brfss). The original NHIS data can be accessed at [www.cdc.gov/nchs/nhis.htm](http://www.cdc.gov/nchs/nhis.htm). The Minnesota Population Center released its 'Integrated Health Interview Series', available at [www.ihis.us/ihis](http://www.ihis.us/ihis), which harmonizes the NHIS data. This study uses the cleaned NHIS data provided by the MPC.

### *Race/Ethnicity Recodes*

All datasets report race/ethnicity in different ways, and the definitions often change over time. The main goal of the recoding was to have consistent samples of the following three groups: Non-Hispanic Whites, non-Hispanic Blacks, and Hispanics. Any observation which did not belong to these three groups, or for which information regarding membership of these three groups was missing, was removed from the sample prior to analysis.

In the NHANES1, 'White' was defined as 'Race = White' and 'Ancestry = German, Irish, Italian, French, Polish, Russian, English, or American'. 'Black' was defined as 'Race = Negro' and 'Ancestry = Negro'. 'Hispanic' was defined as 'Race = White' and 'Ancestry = Spanish or Mexican'.

In the NHANES2, 'White' was defined as 'Race = White' and 'Ancestry = Other European [i.e. excluding Spanish]'. 'Black' was defined as 'Race = Black' and 'Ancestry = Black, Negro, or Afro-American'. 'Hispanic' was defined as 'Race = White' and 'Ancestry = Countries of Central or South America, Chicano, Cuban, Mexican, Mexicano, Mexican-American, Puerto Rican, or Other Spanish'.

In the NHANES3, 'White' was defined as 'Race = White' and 'Ethnicity = Not Hispanic'. 'Black' was defined as 'Race = Black' and 'Ethnicity = Not Hispanic'. 'Hispanic' was defined as 'Ethnicity = Mexican-American or Other Hispanic'. [This was preferred to the classification given by the 'Race-ethnicity' variable because the latter restricts Hispanics to Mexican-Americans.]

In the NHANES4–7, 'White' was defined as 'Race/Ethnicity – Recode [RIDRETH1] = Non-Hispanic White'. 'Black' was defined as 'Race/Ethnicity – Recode = Non-Hispanic Black'. 'Hispanic' was defined as 'Race/Ethnicity – Recode = Mexican American or Other Hispanic'.

In the NHIS, 'Hispanic' was defined as 'Hispanic Ethnicity = Mexican, Mexican-Mexicano, Mexicano, Mexican-American, Chicano, Puerto Rican, Cuban/Cuban American, Dominican (Republic), Other Hispanic, Central or South American, Other Latin American, type not specified, Other Spanish; Hispanic/Latino/Spanish, non-specific type; Hispanic/Latino/Spanish, type refused; Hispanic/Latino/Spanish, type not ascertained; Hispanic/Spanish, type don't know; Multiple Hispanic'. 'White' was defined as 'Race = White' and 'Hispanic = 0'. 'Black' was defined as 'Race = Black/African-American' and 'Hispanic = 0'.

In the BRFSS, 'RACE' was used for 1987–2000 ['New race code', Race/Ethnicity in Eight Groups'], and 'RACE2' from 2001–2008 ['Race/Ethnicity Combination Code', 'Race-Ethnicity grouping', 'Computed Race-Ethnicity grouping']. 'White' was defined as 'Race = White, Non-Hispanic'. 'Black' was defined as 'Race = Black, Non-Hispanic'. 'Hispanic' was defined as 'Race = Hispanic, White; Hispanic, Black; or Other Hispanic' from 1987–2000, and as 'Race = Hispanic' from 2001–2008.

### *Education Recodes*

Reporting of educational attainment also differs greatly between datasets and over time. The different variables were recoded into three measures: (1) A set of three categorical variables that indicate low education (less than 12 years of completed schooling, or no high school degree), medium education (12 years of schooling completed, HS degree, or GED), and high education (more than 12 years, Technical School or College Degrees); (2) A set of five categorical variables that indicate less than 9th grade; 9th – 11th grade, and 12th grade without a diploma; High School Graduation or GED receipt; 1 to 3 years of College or Technical School, or AA degree; 4 or more years of College, or College Graduation; (3) A continuous measure of years of completed schooling, where the average number of years for each of the five categories listed above was taken from the March 2000 CPS: 5.5 years for category 1; 10.3 years for category 2; 12 years for category 3; 14 years for category 4; and 16.6 years for category 5.

In the IHIS version of the NHIS, starting in 1997, it appears that people who finished 12 years of high school but received no diploma were coded as having completed the 12th grade, in contradiction to what was done in previous years. Therefore, for years past 1996, the original NHIS coding was used in this study, where these individuals are counted as having completed only 11th grade.

### *Income Recodes*

All datasets report income in brackets and top-coded. The variables describe either family income (NHIS; NHANES1–3) or household income (BRFSS; NHANES4–7). Several steps were necessary to convert these variables into continuous inflation-adjusted measures of personal income. First, mid-points were assigned to all brackets, and mean values for incomes above the top-codes were imputed from the March CPS files for each year. Second, a measure of personal income was calculated by dividing the family or household income measure by the number of adults per family or household, respectively. Lastly, these personal nominal income measures were inflation-adjusted using the CPI (<ftp://ftp.bls.gov/pub/special.requests/cpi/cpiiai.txt>).

### *Pregnancies*

Pregnancy status is available in the NHIS starting in 1997, in the BRFSS starting in 1985 (i.e. for all years included in this study), and for all series of the NHANES.

### *NHANES Questionnaire Text*

- About how tall are you without shoes?
- About how much do you weigh without clothes or shoes?

### *NHIS Questionnaire Text*

- How tall are you without shoes?
- How much do you weigh without shoes?

### *BRFSS Questionnaire Text*

- What is your height without shoes?
- What is your current weight? (without shoes, with light indoor clothing on)

**Appendix Table 1.** Marginal effects of schooling and income on BMI, NHANES (exam-based)<sup>a</sup>

	1973			1978			1991			1999–2001			2003–2005		
	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]
<b>All males</b>															
Schooling	0.024		0.026	-0.029		-0.023	<i>-0.088</i>		<i>-0.065</i>	-0.038		0.003	-0.044		-0.051
Income		0.001	-0.001		-0.006	-0.005		<i>-0.016</i>	-0.011		<i>-0.016</i>	<i>-0.016</i>		0.000	0.002
N	3,833	3,833	3,833	4,725	4,725	4,725	5,946	5,946	5,946	3,129	3,129	3,129	3,416	3,416	3,416
R <sup>2</sup>	0.033	0.033	0.033	0.054	0.054	0.054	0.041	0.041	0.043	0.025	0.028	0.028	0.019	0.018	0.019
<b>All females</b>															
Schooling	<i>-0.333</i>		<i>-0.249</i>	<i>-0.316</i>		<i>-0.255</i>	<i>-0.297</i>		<i>-0.196</i>	<i>-0.335</i>		<i>-0.222</i>	<i>-0.336</i>		<i>-0.241</i>
Income		<i>-0.085</i>	<i>-0.057</i>		<i>-0.069</i>	<i>-0.044</i>		<i>-0.060</i>	<i>-0.045</i>		<i>-0.052</i>	<i>-0.040</i>		<i>-0.045</i>	<i>-0.032</i>
N	6,300	6,300	6,300	5,066	5,066	5,066	6,335	6,335	6,335	2,974	2,974	2,974	3,215	3,215	3,215
R <sup>2</sup>	0.100	0.096	0.111	0.094	0.085	0.100	0.064	0.067	0.074	0.040	0.044	0.050	0.036	0.035	0.041

<sup>a</sup>Schooling measured in years. Income measured in thousands of constant year 2000 USD. Cells report the coefficients of schooling and real personal income from pooled regressions of BMI on schooling (column [1]), real personal income (column [2]), or both (column [3]). Covariates include age, age squared, age by year, age squared by year, and year fixed effects. All regressions use the appropriate sample weights; standard errors have been clustered by PSU. Italic entries are statistically significant at 5%.

**Appendix Table 2.** Marginal effects of schooling and income on BMI, NHANES (self-reported)<sup>a</sup>

		1973			1978			1991			1999–2001			2003–2005		
		[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]			
<b>All males</b>																
Schooling		<i>-0.076</i>		<i>-0.066</i>	<i>-0.138</i>		<i>-0.117</i>	-0.065		-0.014	<i>-0.079</i>		<i>-0.088</i>			
Income	N/A		<i>-0.012</i>	<i>-0.007</i>		<i>-0.020</i>	<i>-0.011</i>		<i>-0.020</i>	<i>-0.020</i>		<i>-0.002</i>	0.003			
N		4,725	4,725	4,725	5,946	5,946	5,946	3,129	3,129	3,129	3,416	3,416	3,416			
R <sup>2</sup>		0.056	0.054	0.057	0.044	0.040	0.045	0.023	0.028	0.028	0.019	0.018	0.020			
<b>All females</b>																
Schooling		<i>-0.316</i>		<i>-0.259</i>	<i>-0.285</i>		<i>-0.191</i>	<i>-0.349</i>		<i>-0.257</i>	<i>-0.324</i>		<i>-0.239</i>			
Income	N/A		<i>-0.066</i>	<i>-0.041</i>		<i>-0.057</i>	<i>-0.043</i>		<i>-0.046</i>	<i>-0.032</i>		<i>-0.042</i>	<i>-0.028</i>			
N		5,066	5,066	5,066	6,335	6,335	6,335	2,974	2,974	2,974	3,215	3,215	3,215			
R <sup>2</sup>		0.094	0.082	0.101	0.069	0.072	0.079	0.044	0.043	0.052	0.038	0.036	0.043			

<sup>a</sup>Schooling measured in years. Income measured in thousands of constant year 2000 USD. Cells report the coefficients of schooling and real personal income from pooled regressions of BMI on schooling (column [1]), real personal income (column [2]), or both (column [3]). Covariates include age, age squared, age by year, age squared by year, and year fixed effects. All regressions use the appropriate sample weights; standard errors have been clustered by PSU. Italic entries are statistically significant at 5%.

**Appendix Table 3.** Marginal effects of schooling and income on BMI, BRFSS (self-reported)<sup>a</sup>

	1987–1989			1990–1999			2000–2008		
	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]
<b>All males</b>									
Schooling	<i>-0.092</i>		<i>-0.083</i>	<i>-0.100</i>		<i>-0.087</i>	<i>-0.116</i>		<i>-0.093</i>
Income		<i>-0.008</i>	<i>-0.004</i>		<i>-0.009</i>	<i>-0.005</i>		<i>-0.011</i>	<i>-0.006</i>
N	57,559	57,559	57,559	374,618	374,618	374,618	791,780	791,780	791,780
R <sup>2</sup>	0.051	0.048	0.051	0.045	0.042	0.045	0.032	0.030	0.032
<b>All females</b>									
Schooling	<i>-0.314</i>		<i>-0.224</i>	<i>-0.335</i>		<i>-0.222</i>	<i>-0.394</i>		<i>-0.248</i>
Income		<i>-0.050</i>	<i>-0.036</i>		<i>-0.053</i>	<i>-0.041</i>		<i>-0.053</i>	<i>-0.041</i>
N	69,916	69,916	69,916	466,186	466,186	466,186	1,109,219	1,109,219	1,109,219
R <sup>2</sup>	0.093	0.094	0.106	0.073	0.079	0.088	0.048	0.056	0.065

<sup>a</sup>Schooling measured in years. Income measured in thousands of constant USD 2,000---. Cells report the coefficients of schooling and real personal income from pooled regressions of BMI on schooling (column [1]), real personal income (column [2]), or both (column [3]). Covariates include age, age squared, age by year, age squared by year, and year fixed effects. All regressions use the appropriate sample weights; standard errors have been clustered by PSU. Italic entries are statistically significant at 5%.