Income Differentials on Body Mass Index in the U.S. (1971-2002):

A Complex Pattern of Consistency and Change within Race-sex Groups

Virginia W. Chang, MD, PhD

Departments of Medicine and Sociology
University of Pennsylvania

Diane S. Lauderdale, PhD

Department of Health Studies
University of Chicago

[EXTENDED ABSTRACT]

September 23, 2004

Obesity is now described as epidemic and considered to be a major public health concern in the U.S. Recent studies show that the prevalence of obesity has increased from 12.0% to 19.8% over the last decade (Mokdad et al. 2001), and that over 60% the adult population is now overweight or obese (Hedley et al. 2004). While it is frequently noted in the medical and public health literatures that poverty is a risk factor for obesity (in affluent nations), less attention is devoted to the complex relationship between weigh status and socioeconomic status (SES) as a broader gradient. In a review of studies published before 1990, Sobal and Stunkard (1989) find that there is a consistently a strong inverse relationship between SES and obesity among women in "developed societies." Among men, studies varied between finding an inverse relationship, a direct relationship, and no relationship. Sobal and Stunkard note, however, that the focus of most of the studies reviewed was not on the relationship between SES and weight status per se. Similarly, data from recent studies focused primarily on estimating the prevalence of obesity in the U.S. show that the unadjusted prevalence of obesity tends to be higher within lower categories of education for the population as a whole(Mokdad et al. 2001; Mokdad et al. 1999).

Recent studies devoted more specifically to examining the relationship between SES and weight status continue to show a strong inverse relationship among women (and often only among white women when race is accounted for) and rather inconsistent findings among men (Dryson et al. 1992; Flegal, Harlan and Landis 1988a; Flegal, Harlan and Landis 1988b; Goodman 1999; Goodman et al. 2003; Gordon-Larsen, Adair and Popkin 2003; Leigh, Fries and H..B 1992; Molarius et al. 2000; Sarlio-Lähteenkorva and Lahelma 1999; Sarlio-Lähteenkorva, Silventoinen and Lahelma 2004; Wang 2001; Wardle, Waller and Jarvis 2002; Zhang and Wang 2003). Many of these studies are restricted to adolescents, some concern populations outside of the U.S., and many rely on rather limited or crude categorizations of SES. While some do employ more detailed data on income, it is typically either modeled as a linear term or as a small set of wide categories, possibly mis-specifying the relationship. Furthermore, few studies have addressed the question of whether and how such differentials may have changed with time in the U.S., changes which may account for diversity in cross-sectional findings. Flegal, Harlan, and Landis (1988a, 1988b) use nationally representative U.S. data to examine secular trends in the relationship of body mass index (BMI) to education and income over the period 1960-80. Among women, they find that an inverse association with education increased over this time, while an inverse association with income decreased. Among men, they find that the association with education changed from being slightly positive to negative, and that a positive association with income remained constant. Both income and education, however, are modeled as 3-category variables, possibly obscuring finer distinctions. To our knowledge, this work has not been extended to more recent data, and it is after 1980 that the distribution of BMI has shifted dramatically.

In this study, we examine income differentials on BMI using a series of comparable, nationally representative samples of U.S. adults, and we consider whether or not differentials have changed over a three decade period from 1971 to 2002. We also consider the varying nature of this association between race-sex groups, and we employ a more flexible modeling strategy than has been used in past studies to more accurately represent the relationship between BMI and income.

METHODS

To examine secular changes in income differentials on BMI, we use four successive "waves" of the National Health Interview and Examination Survey: NHANES I (1971-1974), NHANES II (1976-1980), NHANES III (1988-1994), and NHANES 1999-2002. The NHANES is conducted by the National Center for Health Statistics to obtain data on the health and nutritional status of the U.S. population through in-person home interviews and direct physical examinations. These surveys are designed to provide nationally-representative, cross-sectional estimates at successive points in time. Measured height and weight, used to calculate body mass index (BMI; defined as weight [kg] divided by the square of height [m]) are included among the physical exam components in each wave. Each survey also collects standard demographic information such as age, sex, race, and household income. We restrict our samples to adults (aged 18-65), non-Hispanic whites and blacks, persons who were examined, and women who are not pregnant. The Hispanic population was not oversampled and cannot be identified in the first two surveys. Sample sizes for each race-sex group in each survey are shown in Table 1.

In the NHANES, data on annual household income is collected in the form of multiple intervals of income, and the cut-points and number of categories change from survey to survey. We transform this data into a continuous variable (1) to permit a more flexible and parsimonious modeling strategy that preserves existing data heterogeneity, and (2) to allow standardization between surveys in the face of considerable inflation over the 32 years. In each survey, respondents are assigned an income corresponding to the midpoint of their income interval. For respondents in the last, open-ended category, a Pareto estimate for the median of this category is used (Parker and Fenwick 1983). To render the continuous income variable we generated into a variable that was comparable between surveys, we standardize income based on the estimated mean and standard deviation of income for each corresponding survey wave. This produces a continuous income variable in each survey with mean zero and standard deviation one. We use income rather than education for two primary reasons. First, the social significance and return for categories of education (e.g., high school graduate) have changed markedly over time. Second, only a very coarse three-level education variable has been released for the most recent NHANES (1999-2002): less than a high school graduate, high school graduate and more than high school graduate. Given the current education distribution in the country, this is not an adequately detailed categorization to address our study questions.

For descriptive statistics, we use the SVY suite of commands in STATA 8 that allow us to account for the NHANES sampling and survey design elements. To model the relationship between BMI and income, we first carried out preliminary analyses using multivariate linear regression to determine the nature of the relationship between age and BMI (as the dependent variable). These analyses revealed a curvilinear relationship with BMI that was well-modeled with the inclusion of a squared term for age. Preliminary analyses also revealed a non-linear relationship between income and BMI, particularly among men. As this concerns our relationship of interest, we fit models using fractional polynomial functions of income to best represent and capture the actual form of the relationship between BMI and income. All models are fit with two degrees for income as higher degrees did not significantly improve model fit. This permits maximal flexibility in modeling while retaining model parsimony and making full use of income heterogeneity in the data. For example, using dummy variables for survey income intervals would require the addition of over 25 terms for some years. Grouping income intervals together, on the other hand, would obscure income heterogeneity that is offered by the data.

The "fracpoly" procedure in STATA 8 was used to select the best-fitting two fractional polynomial terms to capture the relationship between the standardized income variable and BMI for each of the four race-sex groups in each of the four surveys (16 models). This procedure, for example, could select as the best fitting two terms x^2 and ln(x) for one model, while selecting $x^{-.5}$ and $x^{.5}$ for a different model. Fractional polynomial models were adjusted for survey weights

RESULTS

Sample characteristics for each of the four surveys are given in Table 1. The actual numbers of participants meeting study criteria have varied considerably across surveys, as has the extent of oversampling of the black population. Variation in the oversampling greatly affects the precision of estimates for black subsamples, such that NHANES III yields the most precise estimates and consequently the most power to detect significant associations. In NHANES I and II, the extent of oversampling was modest. Average age (among those aged 18 to 64) is generally between 36 and 40 with some evidence of an upward draft over the 32 year period. Income increases approximately six-fold over 32 years. Within race groups, men consistently have higher household income than women. The within-sex ratios of average black income to average white income average about 0.65 without clear trend.

Examining the distribution of each race-sex group into standard clinical categories of BMI shows large differences by race and sex (Table 1). There is modest change in the distributions between the first two surveys, and then evidence of trends from NHANES II through NHANES 1999-2002. The proportion underweight declined, while the proportion in the obese and morbidly obese categories increased. A sex difference in distribution across the categories is less apparent for whites than blacks. For whites, across all surveys, a higher proportion of men than women are in the overweight category, while in the two most recent surveys, a higher proportion of women are in the morbidly obese category. For blacks, there are consistently larger proportions of women than men in both the obese and morbidly obese categories.

Table 2 and the Figures present the results of the fractional polynomial models. Table 2 shows which powers of standardized income were selected for each model and the statistical significance of the two income terms together. At the conventional 0.05 alpha level, the income terms do not achieve statistical significance in three of the sixteen models: black women in NHANES II (p=0.054), white men in NHANES 1999-2002 (p=0.064), and black men in NHANES 1999-2002 (p=0.112). While this might suggest a weakening income effect for men in the most recent years, all three of these p-values are close to significance and in part reflect variations in sample size. Were income modeled as a single linear term, income would appear to be not significantly associated with BMI in several of these models. For example, the p-values for the single linear income term in NHANES II are 0.064 for white men and 0.152 for black men.

The figures graph the predicted BMI values at age forty from each of the 16 fractional polynomial models. They are arranged by race-sex group, showing all four surveys for each group in one graph, allowing one to observe changes in BMI and changes in the shape of the relationship between income and BMI. Note that the y-axis ranges from BMI of 24 to 30 for all groups except Black women, for whom the range is 26 to 32. For all the groups, the curves reflect the distribution changes described in Table 1 with generally similar curves for NHANES I and II, and increasingly high BMI values across the entire income spectrum in the later two surveys.

White women show the most consistent association with roughly parallel curves in all four surveys. There is a clear negative association between income and BMI at all income levels. In NHANES 1999-2002, there appears to be an increasingly strong association in the bottom half of the income distribution, with women below the mean having average BMIs above 28. Note that women at the top of the income distribution in 1999-2002 actually have higher mean BMI than women in the bottom of the income distribution in NHANES I and II. Black women also show an inverse association between income and BMI, and the curves appear to be quite similar to white women in NHANES I. In the more recent surveys, however, the association is flatter for Black women. Lastly, despite these inverse relationships, black women at the top of the income distribution in 1999-2002 have higher mean BMI than women in the bottom of the income distribution in all three previous surveys.

Although generally statistically significant, the patterns are more complicated for men. In NHANES I and II, the association is very similar for black and white men: rising BMI with increasing income, particularly at the low end of the income distribution. Among men, this positive association at the lower incomes continues through all of the surveys for both groups. For white men, however, the inverse association observed for women appears at higher incomes in NHANES II and 1999-2002. For Black men, this inverse association at the higher end of the income distribution does not appear until the most recent survey.

PRELIMINARY CONCLUSIONS

These findings suggest that previous generalizations about the association between income and weight status may have been hampered by a mis-specification of the relationship, particularly among men and blacks, where it is often concluded that there is no association. A complex curvilinear relationship may appear non-significant when modeled as a simple linear relationship or modeled with indicator variables for a limited number of categories. Our analyses also improve on many previous studies by utilizing measured rather than self-reported height and weight. Self-reports are well known to be biased at the extremes of weight, and bias in reporting may be correlated with socioeconomic status as well. The data are also large, nationally-representative samples with the oversampling of the black population, which permits within-race analyses.

[Limitations, more detailed comparisons with prior studies, and estimates of the slope (or change in BMI) at specific incomes will be forthcoming]

REFERENCES

Dryson, E., P. Metcalf, J. Baker, and R. Scragg. 1992. "The Relationship between Body Mass Index and Socioeconomic Status in New Zealand: Ethnic and Occupational Factors." *New Zealand Journal of Medicine* 105:233-235.

Flegal, Katherine M., W.R. Harlan, and J.R. Landis. 1988a. "Secular Trends in Body Mass Index and Skinfold Thickness with Socioeconomic Factors in Young Adult Men." *American Journal of Clinical Nutrition* 48.

——. 1988b. "Secular Trends in Body Mass Index and Skinfold Thickness with Socioeconomic Factors in Young Adult Women." *American Journal of Clinical Nutrition* 48:535-543.

Goodman, Elizabeth. 1999. "The Role of Socioeconomic Status Gradients in Explaining Differences in Us Adolescents' Health." *American Journal of Public Health* 89:1522-1528.

Goodman, Elizabeth, Nancy E. Adler, Stephen R. Daniels, John A. Morrison, Gail B. Slap, and Lawrence M. Dolan. 2003. "Impact of Objective and Subjective Social Status on Obesity in a Biracial Cohort of Adolescents." *Obesity Research* 11:1018-1026.

Gordon-Larsen, Penny, Linda S. Adair, and Barry M. Popkin. 2003. "The Relationship of Ethnicity, Socioeconomic Factors, and Overweight in Us Adolescents." *Obesity Research* 22:121-129.

Hedley, Allison A., Cynthia L. Ogden, Clifford L. Johnson, Margaret D. Carroll, Lester R. Curtin, and Katherine M. Flegal. 2004. "Prevalence of Overweight and Obesity among Us Children, Adolescents, and Adults, 1999-2002." *JAMA* 291:2847-2850.

Leigh, J.P., J.F. Fries, and Hubert H..B. 1992. "Gender and Race Difference in the Correlation between Body Mass and Education in the 1971-1975 Nhanes I." *Journal of Epidemiology and Community Health* 46:191-196.

Mokdad, Ali H., Barbara A. Bowman, Earl S. Ford, Frank Vinicor, James S. Marks, and Jeffrey P. Koplan. 2001. "The Continuing Epidemics of Obesity and Diabetes in the United States." *JAMA* 286:1195-1200.

Mokdad, Ali H., Mary K. Serdula, William H. Dietz, Barbara A. Bowman, James S. Marks, and Jeffrey P. Koplan. 1999. "The Spread of the Obesity Epidemic in the United States, 1991-1998." *JAMA* 282:1519-1522.

Molarius, Anu, Jacob C. Seidell, Susana Sans, Jaakko Tuomilehto, and Kari Kuulasma. 2000. "Educational Level, Relative Body Weight, and Changes in Their Association over 10 Years: An International Perspective from the Who Monica Project." *American Journal of Public Health* 90:1260-1268.

Parker, Robert Nash, and Rudy Fenwick. 1983. "The Pareto Curve and Its Utility for Open-Ended Income Distributions in Survey Research." *Social Forces* 61:872-885.

Sarlio-Lähteenkorva, Sirpa, and Eero Lahelma. 1999. "The Association of Body Mass Index Wtih Social and Economic Disadvantages in Women and Men." *International Journal of Epidemiology* 28:445-449.

Sarlio-Lähteenkorva, Sirpa, Karri Silventoinen, and Eero Lahelma. 2004. "Relative Weight and Income at Different Levels of Socioeconomic Status." *American Journal of Public Health* 94:468-472.

Sobal, Jeffrey, and Albert J. Stunkard. 1989. "Socioeconomic Status and Obesity: A Review of the Literature." *Psychological Bulletin* 105:260-275.

Wang, Youfa. 2001. "Cross-National Comparison of Childhood Obesity: The Epidemic and the Relationship between Obesity and Socioeconomic Status." *International Journal of Epidemiology* 30:1129-1136.

Wardle, Jane, Jo Waller, and Martin J. Jarvis. 2002. "Sex Differences in the Association of Socioeconomic Status with Obesity." *American Journal of Public Health* 92:1299-1304.

Zhang, Qi, and Youfa Wang. 2003. "Socioeconomic Inequality of Obesity in the United States: Do Gender, Age and Ethnicity Matter?" *Social Science and Medicine* 58:1171-1180.

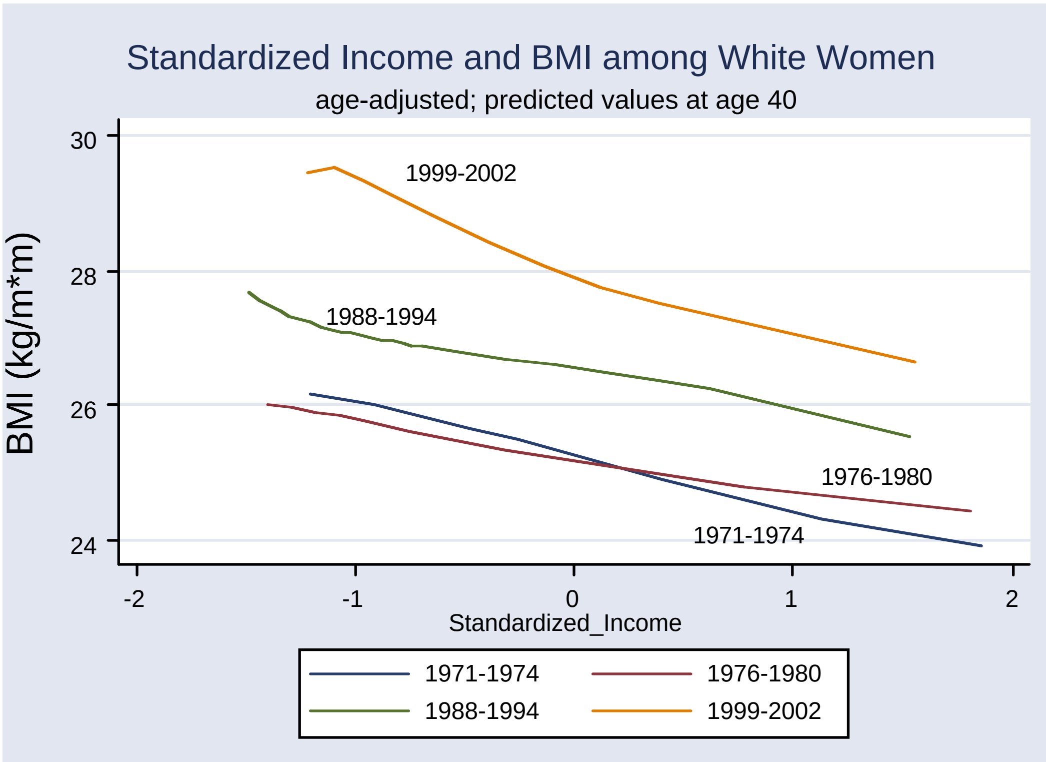
Table 1. Sample characteristics for four surveys, adults aged 18-64. All figures except "Actual number" are adjusted for survey design. BMI categories defined in the text.

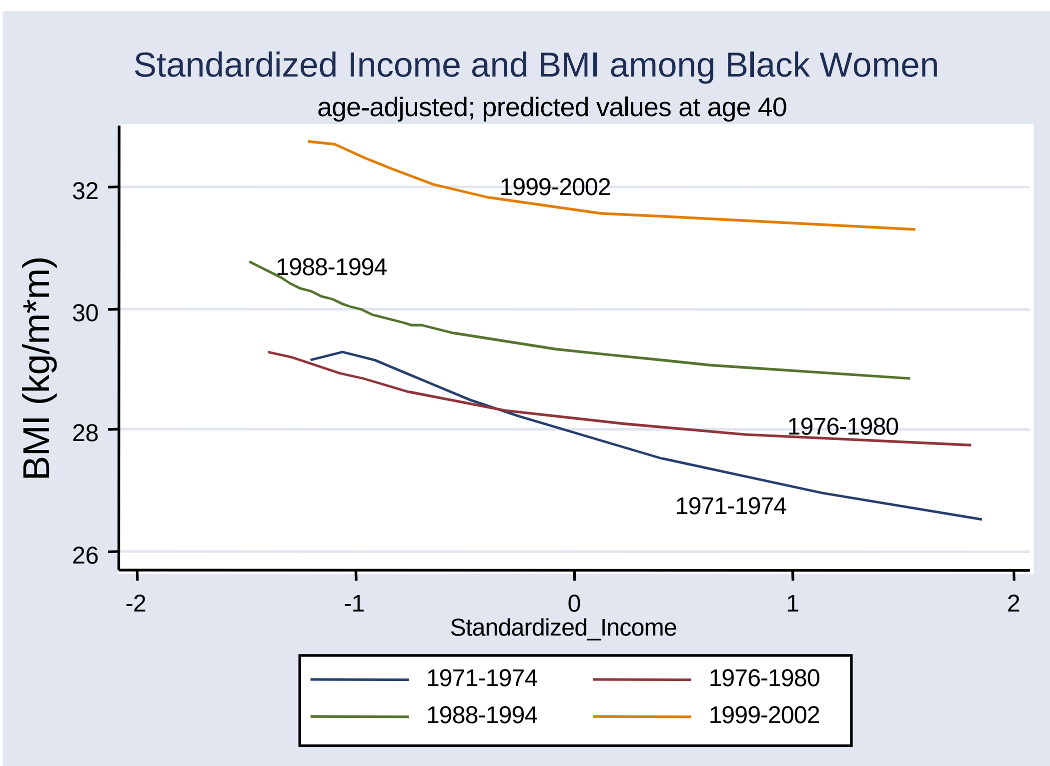
	White Women	Black Women	White Men	Black Men
NHANES I (1971-1974)				
Actual number	6,682	1,375	4,261	683
Population Proportion	0.49	0.10	0.32	0.05
Mean Age	38.4	36.9	41.2	40.3
Mean Income	11,214	6,374	11,902	7,677
BMI categories				
Underweight	0.06	0.06	0.02	0.03
Normal	0.58	0.39	0.45	0.49
Overweight	0.22	0.27	0.41	0.33
Obese	0.09	0.17	0.10	0.12
Morbidly Obese	0.05	0.12	0.02	0.04
NHANES II (1976-1980)				
Actual number	4504	638	4240	539
Population Proportion	0.44	0.06	0.42	0.05
Mean Age	38.9	36.9	38.3	36.7
Mean Income	15,515	10,578	17,235	12,799
BMI categories				
Underweight	0.05	0.04	0.02	0.02
Normal	0.59	0.40	0.48	0.52
Overweight	0.22	0.28	0.39	0.32
Obese	0.09	0.17	0.09	0.12
Morbidly Obese	0.05	0.11	0.02	0.03
NHANES III (1988-1994)				
Actual number	2405	2227	2081	1887
Population Proportion	0.36	0.06	0.36	0.05
Mean Age	39.7	37.2	38.9	36.3
Mean Income	38,530	23,969	40,285	26,145
BMI categories				
Underweight	0.04	0.03	0.01	0.02
Normal	0.51	0.33	0.41	0.43
Overweight	0.22	0.28	0.39	0.35
Obese	0.12	0.18	0.14	0.13
Morbidly Obese	0.10	0.18	0.05	0.07
NHANES 1999-2002				
Actual number	1537	842	1646	841
0.05	0.33	0.06	0.34	0.05
Mean Age	41.4	39.7	40.7	38.3
Mean Income	64,207	39,736	66,854	45,933
BMI categories				
Underweight	0.03	0.01	0.01	0.03
Normal	0.41	0.24	0.31	0.38
Overweight	0.25	0.27	0.40	0.33
Obese	0.16	0.21	0.18	0.15
Morbidly Obese	0.15	0.26	0.09	0.11

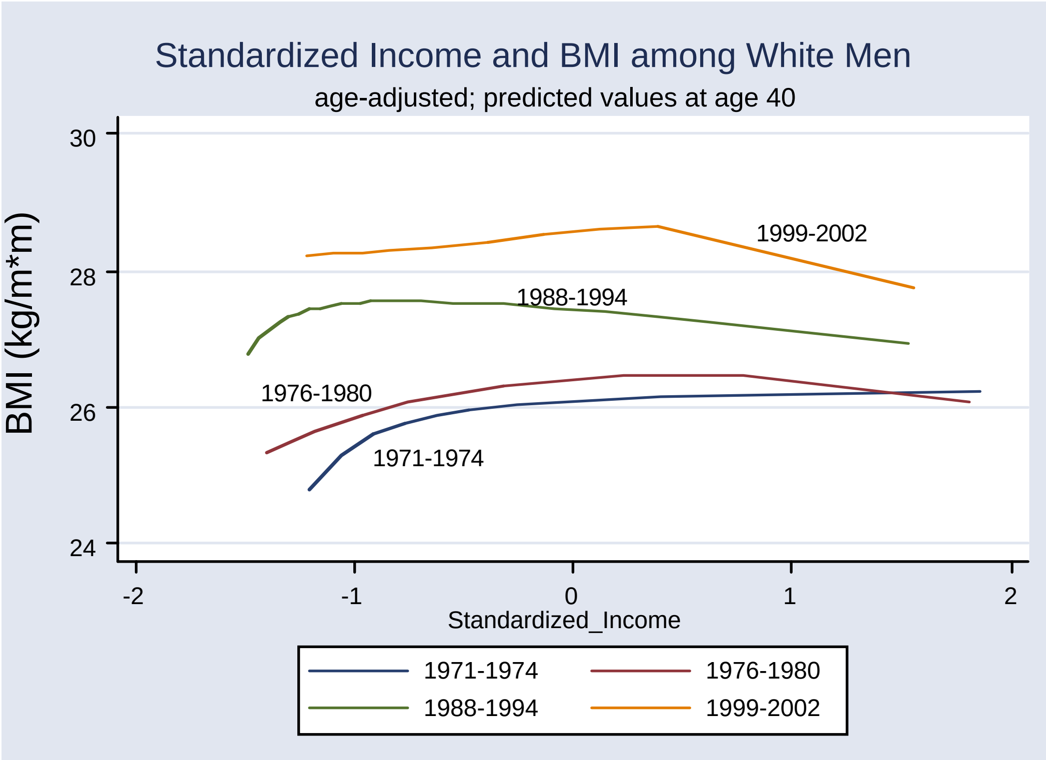
Table 2. From separate fractional polynomial models including two power terms for each race-sex group in each of the four surveys, below are the best-fitting powers of "Standardized Income" (*x*) and a corresponding f-statistic and p-value that test whether the two power terms together are significant additions to the model. Models are also adjusted for age and quadratic age.

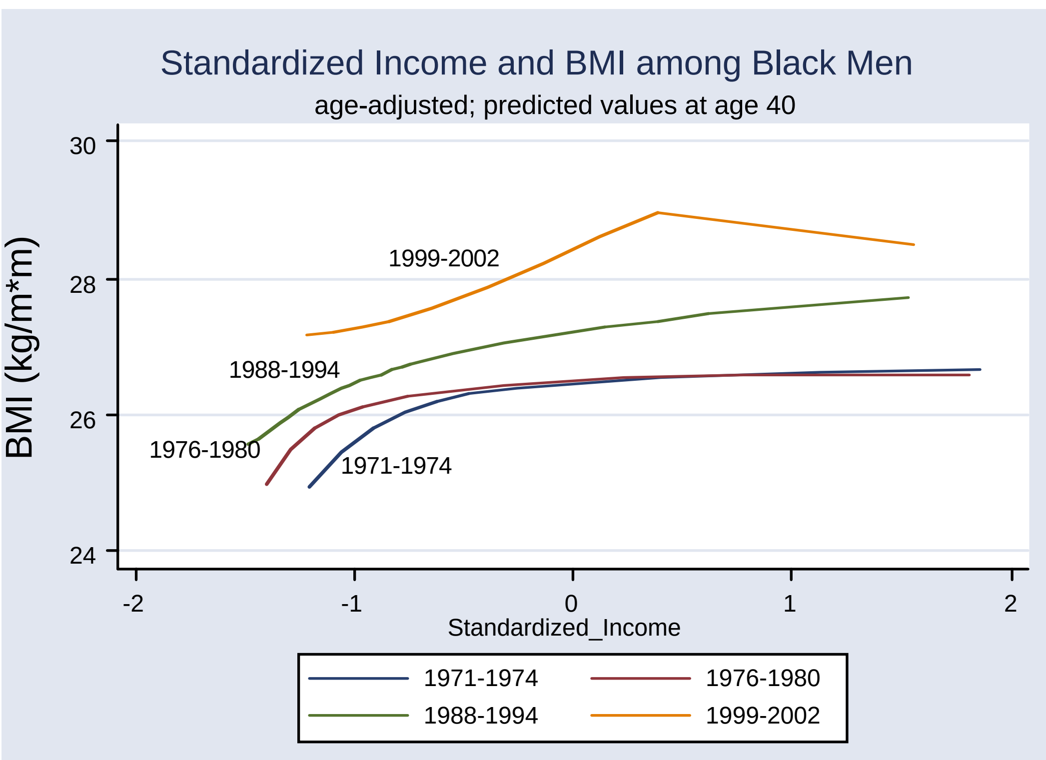
	Best Fitting Powers		**F Stat for both polynomial terms**	**p-value**
NHANES I (1971-1974)				
White Women	x^2	x^2*ln(x)	56.43 (2,6442)	<.0001
Black Women	$x^{-.5}$	$x^{-.5}$*ln(x)	8.08 (2,1311)	.0003
White Men	x^{-2}	X^{-2}*ln(x)	12.04 (2, 4124)	<.0001
Black Men	x^{-2}	X^{-2}*ln(x)	3.22 (2, 646)	.041
NHANES II (1976-1980)				
White Women	$x^{-.5}$	ln(x)	13.54 (2, 4350)	<.0001
Black Women	x^{-1}	X^{-1}*ln(x)	2.93 (2, 599)	0.054
White Men	$x^{.5}$	X^2	12.72(2, 4094)	<.0001
Black Men	x^{-2}	X^{-2}*ln(x)	3.87 (2, 494)	.022
NHANES III (1988-1994)				
White Women	ln(x)	X^3	10.08 (2, 2260)	<.0001
Black Women	$x^{-.5}$	$x^{-.5}$*ln(x)	4.80 (2, 2000)	.008
White Men	$x^{-.5}$	X	5.69 (2, 1977)	.003
Black Men	$x^{-.5}$	$x^{-.5}$*ln(x)	9.21 (2, 1718)	.0001
NHANES 1999-2002				
White Women	$x^{-.5}$	$x^{-.5}$*ln(x)	12.61 (2, 1351)	<.0001
Black Women	x^{-2}	x^{-1}	3.63 (2, 668)	.027
White Men	X^3	X^3*ln(x)	2.75 (2, 1464)	.064
Black Men	X^3	X^3*ln(x)	2.19 (2, 687)	.112

Figures. These figures show the predicted BMI for a forty year old from the fractional polynomial models. Each graph shows predicted values that derive from four separate models, one for each of the four surveys. Models were adjusted for age quadratic age.






Standardized Income and BMI among White Men
age-adjusted; predicted values at age 40
BMI (kg/m*m)
30
28
26
24
1999-2002
1988-1994
1976-1980
1971-1974
-2
-1
0
1
2
Standardized_Income
1971-1974
1976-1980
1988-1994
1999-2002


Standardized Income and BMI among Black Men
age-adjusted; predicted values at age 40
BMI (kg/m*m)
30
28
26
24
1999-2002
1988-1994
1976-1980
1971-1974
-2
-1
0
1
2
Standardized_Income
1971-1974
1976-1980
1988-1994
1999-2002