A Micro-Simulation Based Decomposition of the Health Status Gap Between US Blacks and Whites

Linda Dynan*

ABSTRACT

It is well established that health status differs across racial subpopulations within the United States. Specifically, African Americans (black) live lives that are substantially shorter, on average, than those of their white neighbors. Moreover, blacks generally experience worse health outcomes than whites throughout their lifetimes.

This paper examines the contributions of differences between blacks and whites in specific health-enhancing and health-deterring behaviors to the difference in self-reported health status (and a constructed health status measure) of these two groups. Micro-simulation based decomposition analysis using data from the 2005 Center for Disease Control Behavioral Risk Factor Surveillance System demonstrates that in particular, black/white differences in physical activity have relatively large impacts on the measured health status gap between the two groups, yet black/white differences in socioeconomic and demographic characteristics remain dominant sources in accounting for the observed health status gap.

INTRODUCTION

It is well established that there are underlying differences in health status across racial subpopulations within the United States (Link and Phelan, 1995; Williams and Collins, 1995; Hayward et al., 2000; Institute of Medicine, 2003; Sullivan Commission, 2004; Sequist et al., 2006). Specifically, members of the African-American (black) minority population experience worse health outcomes and live lives that are, on average, substantially shorter than those of their Caucasian (white) neighbors.

In 2001, for example, US life expectancy at birth was 80.2 years for white women and 75.5 years for black women, 75 years for white men and 68.6 years for black men (US Dept. of Commerce, 2004). Although per capita GDP in the US is the fourth highest in the world ($37,562) and per capita health care expenditure is the highest ($5,274 in 2002) (UN, 2005), black men in the United States, on average, live no longer than residents of poor countries with per capita GDPs of approximately $8,000. (Uruguay, for example, in 2002 had per capita income of $8,280, per capita health expenditure of $805, and life expectancy of 75.4 years (UN, 2005).)

---

*Department of Economics and Finance, 426 BEP, Northern Kentucky University, Highland Heights, KY 41099.
◊I would like to thank Rebecca Stein and Kenneth Katkin for their careful reading and thoughtful comments on this work.

Expenditure on health services (intervention and treatment), however, is not sufficient to guarantee good health. In a comprehensive literature review, McGinnis, Williams-Russo and Knickman identify five "domains" that influence an individual's current and future health: genetic and gestational endowment, social circumstances, environmental conditions, behavioral choices, and medical care (McGinnis et al., 2002). Those authors further report:

> On a population basis, using the best available estimates, the impact of various domains on early deaths in the United States distribute roughly as follows: genetic predispositions, about 30 percent; social circumstances, 15 percent; environmental exposures, 5 percent; behavioral patterns, 40 percent; and shortfalls in medical care, 10 percent (McGinnis et al., 2002, p. 83).

Although all of these "domains," and their interactions, are essential to understanding racial variation in health status, this paper focuses in particular on the forty percent component represented by behavioral choices. Focusing on this component may in part develop understanding of the differences in health status between whites and blacks that may be attributable to racial (black/white) differences in participation in health-enhancing behaviors (such as exercise) or health-deterring behaviors (such as smoking). This component also reflects substantial differences for each group in the impact of particular behaviors on overall health status. This understanding of differences in health status (Fairlie, 2003)—difference on average in the participation in specific behaviors and difference in the impact of that behavior on health status—lends itself to an analysis of the health status gap using the decomposition methods first developed by Oaxaca (1973). This decomposition exercise stems from the labor economics literature. Oaxaca decomposition methods have been used to analyze gender and racial differences in wages (for example, Oaxaca, 1973; Blinder, 1973; Oaxaca and Ransom, 1994; Kim and Polachek, 1994; Fairlie and Sundstrom, 1999); and differences in computer ownership and small business ownership (Fairlie, 1999 and 2003). This methodology has also been applied in the health-related literature with respect to race and ethnicity (White-Means, 2000; Wenzlow et al., 2004; Charasse-Pouele and Fournier, 2006). It is particularly useful for identifying and quantifying group differences in measurable characteristics and categorical differences (Fairlie, 2003). In this paper, Oaxaca-Blinder decomposition methods, with recent advances in the technique as applied to nonlinear models developed by Fairlie (2003), are used to examine behavioral sources and their contribution to the white/black health status gap.

Better understanding of the contribution of specific behaviors, and the relative importance of these behaviors to health status, can inform policymakers as they attempt to prioritize among competing policies to narrow the health status gap between US blacks and whites. Such understanding can also inform health care providers and health educators regarding which particular behaviors to emphasize when advising and educating in order to achieve a larger positive impact. In pursuit of this understanding, the following are assessed:

1) Differences in the mean of the probability of health status predicted by the behaviors and characteristics according to racial group, and
2) The contributions of differences in specific behaviors across the two populations to the measured health-status gap.

These population characteristics and behavioral factors are cross-sectionally analyzed using data from the 2005 Center for Disease Control (CDC) Behavioral Risk Factor Surveillance System (BRFSS). This paper follows rules for inclusion in the Institute of Medicine's (2003 p. 41) review of unequal treatment. These rules exclude the impact of differential access and patient preferences on (in this case) perceived health status.

This analysis contributes to the literature by examining the contribution of differences in behavior to the black/white health status gap. The findings suggest that there are statistically significant differences in the distribution of characteristics and behaviors that contribute to measured health status. Differences across the two population groups in the "returns," or coefficients of the structural models associated with the observed behaviors and characteristics, are also found. These differences in returns have been characterized in the literature as the "direct effect of race." The measurement of the direct effect of race may include for example, unobserved influences through omitted variable biases, discrimination, and perhaps differential access or benefit from medical interventions. Consequently, it remains difficult to accurately disentangle and then interpret the findings related to the direct effects of race. Thus, the focus will be on the measured contributions of the black/white differences in behavior.

In particular, this analysis finds that differences in levels of physical activity dominate the behavioral contributions to the health status gap between whites and blacks in this data set. These results remain robust even when an alternative (constructed) measure of self-reported health status is analyzed. However, behavioral contributions as a whole remain a relatively small source of the black/white health status gap relative to the contribution of differences between blacks and whites in the distribution of socioeconomic and demographic characteristics.

The paper proceeds with a brief overview of the literature and discussions of the data and the methods of decomposition analysis. Presentation of the empirical results and a discussion of their implications follow. Finally, conclusions, potential policy issues, and areas for future research are considered.

II. HEALTH DISPARITIES LITERATURE REVIEW

A vast literature spanning numerous disciplines including medicine, health services research, sociology, and epidemiology has developed our understanding of health disparity across populations. A brief survey of the literature begins with McKeown's 1979 work that suggested that the role of medicine in producing health is quite limited. In response to McKeown, others have increasingly

sought understanding of how health is achieved, and particularly to understand why health status varies across populations variously defined by: race, income, income inequality, or nation. For example, Sequist et al.'s (2006) work reinforces McKeown's finding that health is not created by medical care alone. In a study that focused on white and black diabetic patients between 1997 and 2001, Sequist and colleagues (2006) found that provision of better health care services for health-disadvantaged groups diminished, but did not eradicate, differences in health status across racial groups. The literature surveyed below identifies essential variables to include in estimating health status and health status differentials.

Link and Phelan (1995) propose that social conditions are fundamental causes of health status differentials. They argue that the racial health status gap persists primarily because socioeconomic resources offer (to those who possess them) access to a wide array of circumstances and environments that provide advantages in the production and maintenance of health. Their hypothesis is dynamic in predicting that advances in health information will be processed and implemented more quickly and fully by those possessing social advantage, thereby exacerbating existing gaps and allowing for persistence in the health status gap over time. Thus, Link and Phelan advocate the inclusion of socioeconomic resources such as income, education, and health insurance into models of health status. Hayward et al. (2000) also support the inclusion of educational attainment, but interpret it further as a marker of early access to resources. These variables are included in the estimation of health status and perform as expected.

Williams and Collins (1995) find that differences between socioeconomic groups in accessibility, utilization, and quality of care, or differences in the benefits derived from medical care, are contributing factors to the widening inequality in health status. However, they further find that the contribution of medical care is not sufficient to explain all of the observed health disparities. Williams and Collins note that:

> European mortality trends…document that a widening of mortality differences between [socioeconomic status] SES groups is partly due to differences in the decline of mortality from conditions amenable to medical intervention. However, the contribution of medical care is limited. The higher SES groups also experienced larger improvements in mortality than did their lower SES counterparts from those causes of death where medical care does not play a major role. (p. 352)

Like Williams and Collins, many authors have found differences across population subgroups in the magnitude of health benefits derived from medical care. Group differences in such benefits have been attributed, for example, to culturally appropriate/inappropriate interaction during medical encounters; race-matching between physicians and patients; or differences by race in patient compliance (for example, IOM, 2003; Sequist et al., 2006; Sullivan Report, 2004). Although important, the contributions of differences in medical care, or the benefits derived from medical care to the persistence of the health status gap, is not the focus of analysis in this paper.

Further complicating the discussion surrounding health disparities has been a debate about whether relative income inequality, rather than low "absolute" income alone, is bad for health. In response to this debate, Mellor and Milyo (2002) tested whether the statistical aggregates generally measured in studies that show income inequality is detrimental to health reflect causality when also controlling for individual income. Mellor and Milyo (2002) find no consistent evidence linking relative income inequality with health status. Accordingly, nominal individual income—rather than any measure of relative income inequality—is included in this study.

Although she does not focus on health status, White-Means (2000) decomposes medical use among the disabled elderly population by race. White-Means finds that differences in demographic characteristics such as wealth or educational level do not fully explain racial (black/white) differences in the use of physicians' services or prescription drugs.

Charasse-Pouele and Fournier (2006) study the impact of direct racial differences (i.e., when individuals with similar characteristics have different health outcomes indicating differing returns to those characteristics from the structural health equations) and indirect racial differences (i.e., when individuals with different characteristics have different health outcomes, after accounting for potential difference in the "returns") on self-reported health status among South Africans using nonlinear methods. They find that the racial health status gap in South Africa is largely attributable to the superior socioeconomic status (indirect racial effects) of whites, but the direct racial impact is complex, and linked to the indirect (socioeconomic) effects.

This paper uses similar methods to the Charasse-Pouele and Fournier (2006) study to explore and quantify the contribution of differences by groups in behavioral, socioeconomic and demographic characteristics (indirect racial effects) associated with health status on the health status gap between US blacks and whites. The purpose of the exercise is to draw attention to the role policy makers could take in reducing the indirect racial effect stemming, in particular, from differences in health-enhancing behaviors through support of appropriate educational interventions, or subsidization of programs targeted at promoting specific health-enhancing behaviors among blacks.

The literature discussed, although clearly not exhaustive, provides direction for the estimation of health status and insight into the potential sources of group differences in health status. The guidance thus provided is incorporated into the analysis that follows.

III. DATA AND METHODS

The data analyzed in this paper were obtained from the Center for Disease Control's (CDC) 2005 Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS, initiated in 1984, is a cross-sectional telephone survey conducted by the state governments with help from the CDC. Data from a random sample of civilian, non-institutionalized adults (people aged 18 or older) in households (one respondent per household) are collected. Phone numbers are randomly selected. Not included in

"households" are vacation homes, group homes (such as fraternities or shelters), or institutions (such as nursing homes). Because the data and the findings from these data pertain only to the adult population who live in households, this sample is not representative of the whole adult US population, and may in particular under-represent black men who are disproportionately incarcerated and in the non-civilian population. However, health issues in those populations are likely to differ from the larger US population and require different policies to address their needs.

Further, non-coverage may be introduced by the telephone survey method employed while non-response may also introduce biases in the data. The CDC reports in the BRFSS documentation (2004) that, "Although overall approximately 95 percent of US households have telephones, coverage ranges from 87 to 98 percent across states and varies for subgroups as well." The CDC technical notes advise that:

> No direct method of compensating for non-telephone coverage is employed by the BRFSS; however, post-stratification weights are used, which may partially correct for any bias caused by non-telephone coverage. These weights adjust for differences in probability of selection and nonresponse, as well as noncoverage, and must be used for deriving representative population-based estimates of risk behavior prevalence (CDC 2005, first page—no page numbering)

These sample selection issues were addressed by using post-stratification weights for all of the subsequent analyses, which address age, sex, and racial biases.

The 2005 BRFSS collected information on 356,112 adults from 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. The variables of particular interest for this analysis are race and health status. BRFSS identifies five categories of race or ethnicity: white, black, American Indian or Native Alaskan, Asian or Pacific Islander, and Hispanic. In this study, however, race or ethnicity is limited to those who self-identified as non-Hispanic white (278,672) and non-Hispanic black (27,735). BRFSS survey respondents reported their health status on a scale from one to five, one being excellent and five being poor.

The use of self-reported health status in an analysis such as this can be problematic if there are systematic differences in the "cutoffs" (that is the level of a characteristic that distinguishes one health level category from another) according to the racial subgroup to which the respondent belongs. To illustrate this issue, Table 1 identifies by race and health status category the mean number of days during the past month that physical health was self-reported by BRFSS respondents as "not good." For the first three categories of health status, the number of days is fairly close across the racial categories. However, on average, it takes more days of "not good" health for a white respondent to report "fair" of "poor" health status that it does for a black respondent. Because whites may systematically rate themselves as being in better health than blacks when using self-reported "cutoff points," it is possible that self-reported health status may overstate the size of the measured health status gap.

Table 1

Variables Used in the Constructed Health Measure

	Mean Number of Days in Past Month that Health was Not Good		BMI Mean	
Self-Reported Health	White	Black	White	Black
Excellent	0.869	1.144	25.03	27.39
Very Good	1.517	1.445	26.78	28.68
Good	3.262	2.882	28.72	29.95
Fair	11.342	8.959	30.13	31.91
Poor	23.77	20.66	30.35	32.81

To check the robustness of the self-reported health status decomposition results, an alternative dependent variable (health) was constructed that imposed the same cutoff points for both races. The measure was constructed using a combination of body mass index (an objective measure) and the number of days in the past month that the respondent's physical health was "not good" as follows: health=1 (excellent) if the respondent is neither overweight nor obese (BMI < 25) and had less than 2 bad health days in the past 30 days; health=2 (very good) if the respondent reported 2 bad health days and is not obese (BMI <30) or less than 2 bad days but is overweight; health =3 (good) if 3 days are in bad health and respondent is not obese; health=4 (fair) if 4-10 out of the past 30 days are in bad health at any BMI *or* the respondent is obese; and health =5 (poor) if health is not good for more than 10 days at any BMI. The cutoffs are based on the physical health days associated with the self-reported health status measure, but the same cutoffs are used regardless of race. When the BMI classification is included in defining health status, the self-reported health status gap understates this constructed health status gap measure.

A number or variables expected to affect the probability that a respondent self-reports a particular health status were used to estimate separate ordered probit models for each race (white and black) and to estimate pooled coefficients based on the literature survey above. The variables available in BRFSS from which the explanatory variables included in the probit models are constructed are presented in Table 2.

A potential endogeneity problem exists with respect to the explanatory variables associated with "physical activity": poor health status may reduce physical activity, while higher levels of physical activity may improve health status. However, differences in physical activity are quite large across the racial groups (see Table 3). To test the robustness of the analysis with respect to physical activity

Table 2
Explanatory Variables

Explanatory Variable Constructed from	Definition	Omitted Case
Partner	1 if married or member of a couple 0 otherwise	
Education	1=no school or only kindergarten 2=grades one through eight 3=grades nine through eleven 4=high school graduate or GED 5=one through three years of college 6=college graduate or more	no schooling
Emotional support*	1=always 2=usually 3=sometimes 4=rarely 5=never	never
Annual Household income	1=less than $10,000 2=$10,000 to less than $15,000 3=$15,000 to less than $20,000 4=$20,000 to less than $25,000 5=$25,000 to less than $35000 6=$35,000 to less than $50,000 7=$50,000 to less than $75,000 8=more than $75,000	<$10,000
Health plan	1 if insured 0 otherwise	
age	in years	
sex	0 if female 1 if male	
Census Region		northeast
northeast	1 if northeast 0 otherwise	
Midwest	1 if midwest 0 otherwise	
West	1 if west 0 otherwise	
South	1 if south 0 otherwise	
Islands	1 if US island 0 otherwise	
Body Mass Index**	4 digit, no decimal	
Current Smoker	1 if yes 0 otherwise	
Heavy Drinker***	1 if yes 0 otherwise	
High Risk****	1 if yes 0 otherwise	
Routine Checkup	1= within the past year 2=between 1 and 2 years ago 3=between 2 and 5 years ago 4=more than 5 years ago 5=never	never

Fruit Index	1= 0 to less than one serving per day 2=one to two servings 3=three or four servings 5= five or more servings	
Physical activity	1=meets moderate and vigorous 2=meets vigorous physical activity 3=meets moderate physical activity 4=insufficient activity to meet moderate or vigorous 5= no moderate or vigorous physical activity.	No moderate or vigorous

*Response to "How often does the respondent get the social and emotional support he or she needs?"

** Note BMI is an explanatory variable in the self-reported health status model, but not in the constructed health variable model.

*** A heavy drinker consumes more than 2 drinks daily if male or 1 drink daily if female ****high-risk indicates if the respondent has ever participated in behavior that elevates risk for HIV/AIDS.

levels, the analysis was conducted using all five possible health status outcomes, then repeated using only those respondents whose health was "good," "very good," or "excellent"; thus eliminating the group where health status may limit potentially health improving physical activity.

Table 3

Share of Respondents by Race and Health Category

Reporting Sufficient or Insufficient Physical Activity

	Insufficient	Sufficient	Insufficient	Sufficient
	Whites		Blacks	
Excellent	38.09	61.90	51.41	45.59
Very Good	44.61	55.34	53.89	46.11
Good	46.22	53.78	56.44	43.57
Fair	55.60	44.40	62.46	37.53
Poor	63.87	36.13	72.75	27.25
Overall Sample	47.93	52.05	59.78	40.22

An ordered probit model estimates the health statuses of each subpopulation using variables constructed from those identified in Table 2 that are believed to influence health status. The white-black health status gap is defined as the difference between the predicted health status of the white population and the predicted health status of the black population. Once having measured the levels of the characteristics associated with health status for each population (white, black), and how changes in these characteristics affect the probability of reporting a particular health status for each population, we can estimate what the health status of the black population would be if the

characteristics or behaviors possessed by the black population were to yield the same returns to health status as they do to the white population (Ehrenberg and Smith, 2003). These estimates can then be decomposed into the share of each health status gap attributable to differences in behaviors and characteristics of the two groups, and the share attributable to differences in the different health status "returns" to those behaviors and characteristics. Both parts of the decomposition reflect the impact of race. The first part reflects the differences in *observed* characteristics and behaviors, while the second part reflects differences in the health generating process between two groups as well as *unobserved* influences such as discrimination and/or omitted variables. The former may be construed as indirect racial effects on health, and the latter a direct effect of race on health. A more formal description of the process, following Fairlie's (2003) extension of the Blinder-Oaxaca method in a nonlinear model (the probit) throughout the discussion, is presented below.

Health status is predicted from the ordered probit as the sum over all of the outcomes (in this case 1, 2, 3, 4 or 5) of the probability of an outcome multiplied by the value of the outcome. The average probability (not the probability of the average) is represented as $\bar{Y}_i$ where i takes on the values w=white and b=black.

The health status gap, $\bar{Y}w - \bar{Y}b$ can be decomposed into:

$$\bar{Y}_w - \bar{Y}_b = [\sum_{i=1}^{nw} F(X_{i,w}\hat{\beta}_w)/n_w - \sum_{i=1}^{nb} F(X_{i,b}\hat{\beta}_w)/n_b] + [\sum_{i=1}^{nb} F(X_{i,b}\hat{\beta}_w)/n_b - \sum_{i=1}^{nb} F(X_{i,b}\hat{\beta}_b)/n_b]$$

where $F(\bar{X}\beta)$ is the cumulative distribution function from the standard normal distribution and n_j is the sample size for race j. The elements in the first bracket represent the part of the racial gap that is due to differences in the distribution of (all of) the X variables. Elements in the second bracket represent differences in the underlying group processes that generate the levels of Y observed, as well as unmeasured and unobserved characteristics and endowments. Given the more ambiguous interpretation of the second bracketed term, the focus of the analysis will be on the elements in the first bracket. The gap can also be measured using the black beta coefficients as weights in the first term and the white distribution X as weights in the second term. These alternative methods of calculating the gap can lead to different estimates (the indexing problem). For this reason, a range for the health status gap using both methods of weighting is reported.

When determining the contribution of specific variables to the health status gap, as is done here to assess the contribution of differences in the distribution of behavioral variable (high-risk for HIV/AIDS activity, level of physical activity, servings of fruits and vegetables, drinking, and smoking), the calculations are more complicated. The basic calculations (accounting again for the indexing problem by reversing the role of the white and black samples) determine the change in the average predicted probability using a ranked matching of the two samples and then replacing the black distribution for the white distribution *only for the variable of interest* while holding the rest of the white variables constant. The equation below is thus equal to the part of the racial gap that is due to the

difference in the distribution of a specific X variable (corresponding to the first bracket in the previous equation):

$$1/n_b \sum_{i=1}^{nb} F(\hat{\alpha}^* + X_{1iw}\hat{\beta}^*_1 + X_{2iw}\hat{\beta}^*_2) - F(\hat{\alpha}^* + X_{1ib}\hat{\beta}^*_1 + X_{2iw}\hat{\beta}^*_2),$$

where X_{2i} represents the set of variables that remain constant.

However, because the sample sizes differ across racial subpopulations, Fairlie (2003) (2006) suggests the following solution. Use pooled (black and white samples combined) coefficient estimates to calculate predicted probabilities for each observation in the sample. (Note that the pooled estimates contain a race dummy that is then left out of the decomposition analysis because of the focus on differences in the group distribution of behaviors). Draw a random sample of the larger racial population to match the sample size of the smaller racial population. Rank each of the racial samples by their predicted probabilities and match them to each other. Perform the replacement decomposition exercise. Because the decomposition results depend on the random sample chosen, the exercise needs to be repeated a large number of times. (Relying on the central limit theorem, a large number of times will be 30). The mean of the estimates of the repeated samples is then calculated to approximate the true value of the decomposition.

IV. EMPIRICAL RESULTS AND DISCUSSION OF IMPLICATIONS

Mean values (or sample proportions) by racial group from the 2005 BRFSS for the dependent variables and the variables included in the ordered probit model are presented in Table 4. (Panel A of Table 4 presents difference in means for continuous and indicator variable; Panel B of Table 4 presents differences in sample proportions for categorical variables).

TABLE 4
PANEL A
Means (and 0/1 Indicator Variables) and Differences in Means
from the 2005 BRFSS

Variable	*White*	*Black*	*W-B*
Age	44.91	42.42	2.50*
			(0.00)
BMI	27.29	29.61	-2.32*
			(0.00)
Sex (Males)	0.414	0.319	0.095*
			(0.00)
Partner	0.660	0.363	0.297*
			(0.00)
Health Plan	0.877	0.807	0.070*
			(0.00)

Exercise	0.801	0.677	0.123*
			(0.00)
Smoker	0.228	0.228	0.000
			(0.94)
Drinker	0.056	0.034	0.022*
			(0.00)
High Risk	0.025	0.055	-0.030*
			(0.00)
Northeast	0.202	0.148	0.055*
			(0.00)
Midwest	0.237	0.169	0.068*
			(0.00)
South	0.207	0.405	-0.197*
			(0.00)
West	0.265	0.045	0.219*
			(0.00)
Islands	0.002	0.054	-0.052*
			(0.00)

P-values are in parentheses.
* indicate significance at <1%

TABLE 4
PANEL B
Sample Proportions and Differences in Sample Proportions from the 2005 BRFSS

Variable	*Health Status (1=Excellent to 5=poor)*			
	White Self Report	*White Constructed*	*Black Self Report*	*Black Constructed*
Excellent	23.73	28.10	16.44	17.81
Very Good	37.94	30.76	28.87	28.11
Good	26.19	2.61	34.34	2.18
Sum	87.86%	61.47%	79.65%	48.10%
Fair	8.44	28.11	15.12	39.37
Poor	3.70	10.42	5.23	12.53
Sum	12.14%	38.53%	20.35%	51.9%
Mean Values	2.30	2.62	2.64	3.00

Pearson Chi2 4.8e+03
P (0.00)

	White	*Black*	*W-B*
Checkup			
Never	1.13	0.41	0.72
Past year	64.56	78.81	-14.25
1-2 years	15.07	11.46	3.61
2-5 years	9.48	5.51	3.97
5+ years	9.77	3.82	5.95

Pearson Chi2 2.3e+03
P (0.00)

Education			
None	0.04	0.04	0
grades 1-8	0.83	1.79	-0.96
grades 9-11	4.0	8.96	-4.96
HS/GED	27.03	34.20	-7.17
some college	28.34	29.18	-0.84
college grad or more	39.75	25.83	13.92
	Pearson Chi2 1.3e+04		
	P (0.00)		
Income			
>$10K	3.54	10.32	-6.78
$10K-<$15K	3.53	7.28	-3.75
$15K-<$20K	4.76	17.12	-12.36
$20K-<$25K	7.29	11.97	-4.68
$25K-<$35K	11.74	15.82	-4.08
$35K-<$50K	18.15	16.65	1.5
$50K-<$75K	21.37	13.03	8.34
> $75K	29.61	12.86	16.75
	Pearson Chi2 1.4e+04		
	P (0.00)		
Variable	*White*	*Black*	*W-B*
Emotional Support			
Always	45.05	45.18	-0.03
Usually	36.90	22.15	14.52
Sometimes	12.33	21.11	-8.78
Rarely	3.81	5.84	-2.03
Never	1.91	5.72	-3.81
	Pearson Chi2 4.7e+03		
	P (0.00)		
Fruits and Vegetables			
>1serving	4.24	6.62	-2.38
1-2servings	35.60	37.57	-1.97
3-4 servings	36.58	32.71	3.87
5+ servings	23.58	23.09	0.49
	Pearson Chi2 557.67		
	P (0.00)		
Physical Activity Level			
Moderate and Vigorous	18.00	10.31	7.69
Vigorous	11.46	12.66	-1.2
Moderate	22.52	17.15	5.37
insufficient	38.62	40.93	-2.31
none	9.40	18.96	-9.56
	Pearson Chi2 3.3e+03		
	P (0.00)		

From Table 4, it is clear that there are statistically significant differences in the mean values of the identified characteristics and behaviors across the populations examined. The statistical significance of the differences is practically uniform for the characteristics and behaviors under study, with the exception that the proportion of smokers in the white population is not statistically significantly different from the proportion of smokers in the black population. All of the other variables are significantly different across populations at levels of significance up to five percent. The mean health status gap between whites and blacks is 0.332203 in favor of whites (about 14.4 percent when divided by predicted white health status: 2.304).

The underlying structural equations of the health probability function also differ across the two racial groups. Because the variable to be estimated is an ordered qualitative variable, an ordered probit was used to estimate health status based on the other variables thought to underlie the respondent's self-reported (and constructed) health status. The results of the three (white, black and pooled) ordered probits are reported in Appendix Table A1 (self-reported health status) and Table A2 (constructed health status). The data are weighted using post-stratification weights to correct for non-response biases in the sample, and non-coverage of households without telephone services along race, sex, and age dimensions. Usable observation rates are 58 percent for whites and 60 percent for blacks.

It is important to remember, however, that the predicted health status here is merely an estimate of a probability function. The coefficients cannot be interpreted as marginal effects. The marginal impact of these variables on the probability of reporting a particular health status is calculated from the result of the ordered probit. The marginal impacts for each of the explanatory variables are available from the author.

Using the first decomposition method described above (and the reverse ordering to account for the indexing problem),

$$\bar{Y}_w - \bar{Y}_b = [\sum_{i=1}^{nw} F(X_{i,w}\hat{\beta}_w)/n_w - \sum_{i=1}^{nb} F(X_{i,b}\hat{\beta}_w)/n_b] + [\sum_{i=1}^{nb} F(X_{i,b}\hat{\beta}_w)/n_b - \sum_{i=1}^{nb} F(X_{i,b}\hat{\beta}_b)/n_b]$$

to estimate the health status gap between blacks and whites of the entire set of independent variables finds (see Table 5): the white-black gap is 0.332203 (favoring whites) and the contribution from group specific differences in the distribution of the independent variables ranges from 61.88 percent to 95.99 percent When the constructed health variable is used, the gap is 0.378498 favoring whites (a gap of about 14.6 percent=0.37849/2.62), and the contribution from group specific differences in the distribution of the independent variables ranges from 26.46 percent to 58.34 percent. This suggests that if all the measured characteristics and behaviors of blacks and whites (behavioral characteristics as well as socioeconomics and demographic characteristics) were identical, the health status gap between them would decline from 14.4 percent to between 0.7 percent and 5.6 percent (using self-reported values) but still not disappear. The same measures using the constructed measure show a

similar decline, but failure to close, in the gap: from 14.6 percent to between 5.3 percent and 9.3 percent.

TABLE 5

All Variables Decomposition

Dependent Variable: Self-Reported Health Measure

Order	Gap	Group difference Distribution	Contribution Group Difference Distribution
White-Black	-0.332203	-0.318872	95.99%
Black-White	-0.332203	-0.205558	61.88%

Range of contribution of Group Differences in Distribution: 61.88%-95.99%

Dependent Variable: Constructed Health Measure

Order	Gap	Group difference Distribution	Contribution Group Difference Distribution
White-Black	-0.378498	-0.220821	58.34%
Black-White	-0.378498	-0.100151	26.46%

Range of contribution of Group Differences in Distribution: 26.46%-58.34%

Specific Variable Contributions: Self-Reported Dependent Variable

The results for the decomposition exercise identifying the contribution of differences across racial groups in the distribution of specific individual behaviors to the health status gap are presented in Table 6. Standard errors are calculated according to the delta method proposed by Fairlie (2003). Ranges are presented to account for the indexing problem noted above. These calculations of the health status gap use pooled coefficients with a racial dummy included in estimating the coefficients, but the racial dummy is then set equal to zero to assess the impact of difference by race in the distributions of characteristics and behaviors. The self-reported health status gap due to differences in distribution of the full set of characteristics using the pooled estimation is 0.304492 (91.66 percent of the gap), a figure within the range reported in Table 5 (61.88 percent to 95.99 percent).

The findings are as follows: black/white differences in the share of each subpopulation that has engaged in high risk behaviors contributes between 0.56 percent and 0.57 percent of the overall health status difference; physical activity contributes between 14.67 percent and 15.93 percent; diet contributes between 0.66 percent and 0.68 percent. Smoking differences yielded no significant differences by race (and as noted, the difference in means by race for smoking was also not significant as documented in Tables 4). Drinking behavior was not significant in the probit model, and although

TABLE 6

Specific Behavioral Variable Contribution
To Health Status Gap
Pooled Coefficient Method

Dependent Variable: Self-Reported

Pooled Estimates: White Health=2.299816
Black Health (race dummy=0)=2.604308
Distributional Gap=-0.304492/0.332203= 91.66%

	B Replaces W	W Replace B
High Risk	0.00186	0.0018894
	(0.00021)	(0.00022)
	0.56%	0.57%
Physical Activity (5 outcomes)	0.0529079	0.0487349
	(0.00038)	(0.00038)
	15.93%	14.67%
Physical Activity (3 outcomes--	0.0209485	0.0211616
gap=-0.139727/0.308333=45.32%)	(0.00049)	(0.0005)
	6.79%	6.86%
Fruits and Vegetables	0.0022899	0.00218431
	(0.0001)	(0.0001)
	0.69%	0.66%

Dependent Variable: Constructed

Pooled Estimates: White Health=2.629956
Black Health (race dummy=0)=2.835241
Distributional Gap=-0.205285/-0.378498= 54.24%

	B Replaces W	W Replaces B
High Risk	0.0033824	0.0033812
	(0.0002)	(0.0002)
	0.89%	0.89%
Physical Activity (5 outcomes)	0.0642087	0.0627651
	(0.0004)	(0.00037)
	16.96%	16.58%
Physical Activity (3 outcomes)	0.006833	0.0824342
Gap=0.009059/-0.087293=-10.37%	(0.0007)	(0.0007)
	7.83%	-94.43%
Fruits and Vegetables	0.000961	0.0009213
	(0.0001)	(0.0001)
	0.25%	0.24%
Drinking	0.0030162	0.0029071
	(0.0001)	(0.0001)
	0.79%	0.77%

Delta method SEs are reported in parentheses.

Dependent Variable: Constructed, No Smokers

Pooled Estimates: White Health=2.597028
Black Health (race dummy=0)=2.782873
Distributional Gap=-0.185845/0.388827=47.80%

	B Replaces W	W Replaces B
High Risk	0.0077996	0.0027536
	(0.0002)	(0.00024)
	2.01%	0.71%
Physical Activity	0.0681445	0.0654424
	(0.0004)	(0.00044)
	17.53%	16.83%
Fruits and Vegetables	0.002427	0.0024061
	(0.0002)	(0.00015)
	0.62%	0.62%
Drinking	0.0029708	0.0029021
	(0.0001)	(0.00015)
	0.76%	0.75%

Delta method SEs are reported in parentheses.

differences by drinking behavior were statistically significant, the magnitude of the contributions from drinking was less than 0.00 percent.

As noted above, differences in distribution of characteristics and behaviors account for approximately 91.66 percent of the health status gap, 15.89 percent-17.18 percent of which is due to behavioral differences and about 74.48 percent to 75.77 percent is due to differences in socioeconomic and demographic characteristics. Thus differences in "returns," or direct effects of race, account for approximately 8.34 percent of the health status gap. Although important, behavioral differences between blacks and whites contribute substantially less than socioeconomic and demographic characteristic differences to the health status gap between blacks and whites.

Because physical activity may influence health and health may influence physical activity, the analysis with respect to physical activity was conducted a second time, including only those respondents who reported good or better health as a sensitivity test. Although the contribution of differences in physical activity to the health status gap diminished in magnitude when respondents who reported less than good health were removed from the data set, these differences nonetheless remained the largest contributor to the health status gap (6.79 percent to 6.86 percent) from the behavioral variables tested.

Overall, differences in physical activity were found to make up the largest contribution to the health status gap when assessing the contribution of differences in the distribution of included behaviors to the racial health status gap based on self-reported health status in this data set. If blacks were to adopt the physical activity behaviors of whites, the health status gap would narrow from 14.4 percent to between 12.12 percent and 12.30 percent as measured from self-reported health status in the unrestricted (all five outcomes) data set.

Specific Variable Contributions: Constructed Dependent Variable

The health status gap using the constructed measure is 0.378498 (about 14.6 percent =

.378498/2.6). Following the same method as that used to conduct the analysis for self-reported health status, the constructed health status gap due to differences in distribution using the pooled estimation method is 0.205285 (54.24 percent of the gap) and is in the range reported in Table 5 (26.46 percent to 58.34 percent).

The findings are as follows: difference by race in the share of the subpopulation who have engaged in high-risk behaviors contributes 0.89 percent of the overall health status difference; physical activity contributes between 16.58 percent and 16.96 percent; diet contributes between 0.24 percent and 0.25 percent; and drinking behavior contributes between 0.77 percent and 0.79 percent. Smoking differences yielded no significant differences by race (the difference in means by race for smoking was also not significant as documented in Tables 4A).

The findings using the constructed health measure reinforce the self-reported health status findings. Again, black/white differences in physical activity are the largest measured behavioral contributor to the black/white health status gap. If blacks were to adopt the same physical activity levels reported by whites, the health status gap between them would narrow from 14.6 percent to between 11.99 percent and 12.05 percent.

The analysis using the constructed estimates of health faced the same endogeneity issues with physical activity as the self-reported measure. Therefore, a sensitivity test was conducted again with the dependent variables restricted to the three better health status outcomes (good, very good or excellent). The measured health status gap between whites and blacks in this group was much smaller (-0.087293) than in the full sample (-0.378498). This result suggests that the distributional differences across racial groups would favor blacks (reducing the gap by 0.009050, to -.078234), and is explained not by differences in the distributions of characteristics and behaviors, but rather by the health generating processes and unexplained elements that contribute to health status. However, when the two samples were restricted to “good or better” health, 61.47 percent of the white sample achieved “good or better” health status under the constructed measure while only 48.09 percent of the black sample did. This result suggests that the bulk of the distributional difference in behaviors may be among respondents in “fair” to “poor” health.

When the physical activity replacement exercise was simulated repeatedly in the restricted sample, the findings suggest that if whites had engaged in physical activity only to the same extent as blacks, their health status would worsen by 0.006833 or 7.8 percent of the gap. The health status of blacks if they engaged in physical activity at the same level as whites would improve by 0.0082342, or close about 94 percent of the remaining gap between blacks and whites.

Finally, the coefficient on smoking yielded a sign contrary to expectation in the probit model. Smoking was construed to be negatively associated with BMI that in part determines health status under the constructed measure. The analyses were again conducted using the constructed measure, but only including nonsmokers in the samples. These analyses are presented and, although differing

in magnitudes, the relative importance of differences in physical activity across racial groups in the health status gap is underscored.

V. CONCLUSIONS, POLICY ISSUES, AND AREAS FOR FUTURE RESEARCH

Although the analyses presented here do not establish causation, the descriptive results reported can contribute to efforts to narrow the health status gap because they identify sources, and levels of contribution of the identified sources, to the observed health status gap. This analysis finds that 61.88 percent to 95.99 percent of the measured self-reported health status gap between blacks and whites can be attributed to differences in behaviors and other socioeconomic and demographic characteristics. (If self-reported health status is replaced by a constructed measure of health status based on self-reported days in poor health and BMI, then the percentage of the measured health status gap between blacks and whites that can be attributed to differences in behaviors and other socioeconomic and demographic characteristics is 26.46 percent to 58.34 percent). Yet even if all the measured characteristics and behaviors were identical across the two groups, the health status gap between them would decline to between 0.7 percent and 5.6 percent (using self-reported values) or to between 6.0 percent and 10.6 percent (using the constructed measure), but still not disappear.

Nonetheless, further efforts to understand which behaviors differ, and why, are likely to improve the health status of minority populations by giving rise to policies that promote specific health-enhancing behaviors. Such policies might involve expanding choice sets, or initiating behavioral change through strong promotion by health care providers, public health agencies and in public schools, towards specific health enhancing behaviors through education and advisement.

In particular, a relatively large share of the gap in terms of behavioral differences—ranging from 14.67 percent to 16.96 percent of the overall approximately 14 percent health status gap —was attributable to difference in physical activity in the unrestricted (five health status level, self-reported measure) samples. If the physical activity gap were to close, the health status gap favoring whites has the potential to narrow to between 11.99 percent and 12.30 percent. Policies designed to enhance physical activity among African American would thus seem to be a promising source of immediate action to reduce the health status gap. Such policies might include community recreation centers located in predominantly African American communities, tax waivers for health fitness facilities that locate in predominantly African American communities, or subsidies to Medicaid recipients for fitness facility memberships. Alternatively, efforts to create built communities to enhance physical activity (green spaces and walkability) can be supported and subsidized by government.

Although physical activity had a relatively large impact in the domain of behavior, behavioral aspects were dwarfed by the contribution of the other socioeconomic and demographic attributes to the health status gap consistent with what has been reported in the literature on health disparities. However, the present analysis suggests that emphasis on enhanced physical activity among blacks

would be an appropriate strategy as it promises the relatively largest reduction in the health status between blacks and whites among the behaviors measured in the analysis. Yet, these efforts alone, as has been found in this analysis, will not be sufficient to close the health status gap between blacks and whites. There remain differences in socioeconomic and demographic characteristics of blacks and whites as well as in the underlying health-generating process and unobserved endowments that will need to be measured and addressed.

Research into specific health behaviors by race and ethnicity, sex, and health status will improve society's ability to target specific groups in ways that will help close the health status gap. Further research that investigates differences in improvements to health status for given behavioral investments (such as quitting smoking, increasing exercise—or engaging in other health-generating behavior) and other behavioral differences that were not captured in BRFSS data analyzed in this study (such as caloric intake, compliance with disease screening recommendations) may serve to improve our understanding, and our ability to reduce the health status gap that has persisted between U.S. blacks and whites for too long.

REFERENCES

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *The Journal of Human Resources,* 8(4): 436-455.

Center for Disease Control. 2006. 2005 Overview Behavioral Risk Factor Surveillance System. December. http://www.cdc.gov/brfss/

Charasse-Pouele, Cecile and Martin Fournier. 2006. "Health Disparities Between Racial Groups in South Africa: A Decomposition Analysis." *Social Science and Medicine,* 62: 2897-2914.

Ehrenberg, Ronald G. and Robert S. Smith. 2003. *Modern Labor Economics*, 8th ed. Addison Wesley.

Fairlie, Robert W. 1999. "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment." *Journal of Labor Economics,* 17(1): 80-108.

Fairlie, Robert W. 2003. "An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models." Yale University Economic Growth Center, Discussion Paper Number 873.

Fairlie, Robert W. and William. A. Sundstrom. 1999. "The Emergence, Persistence, and Recent Widening of the Racial Unemployment Gap." *Industrial and Labor Relations Review,* 52(2): 252-270.

Hayward, Mark. D., Toni P. Miles, Eileen M. Crimmons, and Yu Yang. 2000. "The Significance of Socioeconomic Status in Explaining the Racial Gap in Chronic Health Conditions." *American Sociological Review*, 65: 910-930.

Institute of Medicine. 2003 *Unequal Treatment: Confronting Racial and Ethnic Disparities in Healthcare*. Washington, DC: National Academies Press.

Kim, Moon-Kak and Solomon. W. Polachek. 1994. "Panel Estimates of Male-Female Earnings Functions." *Journal of Human Resources,* 29(2): 406-428.

Link, Bruce G. and Jo Phelan. 1995. "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior,* 35(extra issue): 80-94.

McGinnis, J. Michael, Pamela Williams-Russo and James R. Knickman. 2002. "The Case for More Active Policy Attention to Health Promotion." *Health Affairs,* 21(2): 78-93.

McKeown, Thomas. 1979. *The Role of Medicine: Dream, Mirage, or Nemesis?* New Jersey, Princeton UP.

Mellor, Jennifer M. and Jeffrey Milyo. 2002. "Income Inequality and Health Status in the United States: Evidence from the Current Population Survey." *Journal of Human Resources,* 37(3): 510-539.

Oaxaca, Ronald. L. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review*, 14 (3): 693-709.

Oaxaca, Ronald L. and Michael R. Ransom. 1994. "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics,* 61(1): 5-21.

Sequist, Thomas D., Alyce Adams, Fang Zhang, Dennis Ross-Degnan, John Z. Ayanian. 2006. "Effect of Quality Improvement on Racial Disparities in Diabetes Care. *Archives of Internal Medicine,* 166: 675-681.

The Sullivan Commission. 2004. Missing Persons: Minorities in the Health Profession. http://www.aacn.nche.edu/Media/pdf/SullivanReport.pdf

United Nations*. 2005. Human Development Report 2005: International Cooperation at a Crossroads: Aid, Trade and Security in an Unequal World*. http://hdr.undp.org/en/reports/

United States Department of Commerce, various years. *Statistical Abstract of the United States*. Washington, DC, US GPO. http://www.census.gov/prod/www/statistical-abstract-1995_2000.html

Wenzlow, Audra. T., John Mullahy, and Barbara L. Wolfe. 2004. "Understanding Racial Disparities in Health: The Income-Wealth Paradox." University of Wisconsin, Institute for Research on Poverty, Discussion Paper 1283-04. http://www.pophealth.wisc.edu/phs548/wenzlow.pdf

White-Means, Shelley. I. 2000. "Racial Patterns in Disabled Elderly Persons' Use of Medical Services." *Journal of Gerontology,* 55B (2): S76-S89.

Williams, David. R. and Chiquita Collins. 1995. "US Socioeconomic and Racial Differences in Health: Patterns and Explanations." *Annual Review of Sociology,* 21: 349-86.

Appendix

TABLE A1

Dependent Variable=Self-Reported Health Status (1=excellent to 5=poor)
Ordered Probit Coefficients

Variable	*White*	*Black*	*W/B Pooled*
Constant	0.676	1.315	0.704
Black			0.012
			(0.289)
Age	0.0001	-0.002	0.001
	(0.92)	(0.66)	(0.722)
Age^2	0.0001*	0.000*	0.000
	(0.00)	(0.00)	(0.00)
Health Plan	0.001	-0.21	-0.005
	(0.89)	(0.43)	(0.568)
Checkup (never)			
past year	0.116*	0.072**	0.110*
	(0.00)	(0.02)	(0.000)
past 2 years	0.013	-0.074	0.006
	(0.27)	(0.12)	(0.606)
past 5 years	0.034*	0.059	0.035*
	(0.01)	(0.33)	(0.007)
past	–0.103*	-0.210	-0.107*
	(0.00)	(0.23)	(0.00)
Partner	0.071*	0.034	0.061*
	(0.00)	(0.11)	(0.00)
Education (none)			
1-8^{th}	0.438*	-0.325	0.324**
	(0.00)	(0.41)	(0.021)
1-11^{th}	-0.150*	-0.045	-0.152*
	(0.00)	(0.59)	(0.00)
1-12/GED	-0.197*	-0.102*	-0.174*
	(0.00)	(0.01)	(0.00)
1-some college	-0.090*	-0.048**	-0.080*
	(0.00)	(0.05)	(0.00)
1-college +	-0.156*	-0.068*	-0.152*
	(0.00)	(0.01)	(0.00)
Income (low)			
10K or more	-0.073*	-0.190*	-0.080*
	(0.00)	(0.00)	(0.00)
15K or more	-0.201*	-0.018	-0.168*
	(0.00)	(0.70)	(0.00)
20K or more	-0.139*	-0.151*	-0.134*
	(0.00)	(0.00)	(0.00)
25K or more	-0.164*	-0.059***	-0.145*
	(0.00)	(0.10)	(0.00)
35K or more	-0.115*	-0.113*	-0.115*
	(0.00)	(0.00)	(0.00)
50K or more	–0.088*	-0.088*	-0.090*
	(0.00)	(0.01)	(0.00)
High Income	-0.145*	-0.154*	-0.150*
	(0.00)	(0.00)	(0.00)
Sex	0.038*	-0.030	0.034*
	(0.00)	(0.16)	(0.00)

Emotional Support (none)			
Rarely	0.284*	0.300*	0.301*
	(0.00)	(0.00)	(0.00)
Sometimes	0.075*	0.102**	0.094*
	(0.00)	(0.03)	(0.00)
Usually	-0.161*	0.006	-0.120*
	(0.00)	(0.91)	(0.00)
Always	-0.363*	-0.176*	-0.314*
	(0.00)	(0.00)	(0.00)
Smoker	0.328*	0.182*	0.315*
	(0.00)	(0.00)	(0.00)
Drinker	0.004	-0.027	-0.002
	(0.77)	(0.62)	(0.871)
BMI	0.0004*	0.0003 *	0.0004*
	(0.00)	(0.00)	(0.00)
Fruits/Veg	-0.036*	-0.038*	-0.037*
	(0.00)	(0.00)	(0.00)
Exercise (none)			
Insufficient	-0.368*	-0.249*	-0.341*
	(0.00)	(0.00)	(0.00)
Moderate	-0.458*	-0.253*	-0.425*
	(0.00)	(0.00)	(0.00)
Vigorous	-0.632*	-0.461	-0.597*
	(0.00)	(0.00)	(0.00)
Mod+Vig	-0.706*	-0.574*	-0.680*
	(0.00)	(0.00)	(0.00)
High Risk	0.088*	0.075***	0.073*
	(0.00)	(0.08)	(0.00)
Midwest	0.001	0.004	-0.001
	(0.93)	(0.88)	(0.944)
South	0.033*	-0.012	0.027*
	(0.00)	(0.57)	(0.001)
West	0.031*	-0.006	0.025*
	(0.00)	(0.90)	(0.001)
Islands	-0.125***	0.034	0.043
	(0.07)	(0.38)	(0.206)
Predicted Health Status	2.304044	2.636247	2.33551
Sample Size	161,794	16,780	178,574
Wald Chi2	35799.71	3033.66	37643.42
Prob >chi	0.00	0.00	0.00

Predicted health status is calculated as:
[1-Φ(βhatXbar)]*0 + [Φ(μ1- βhatXbar) - Φ(-βhatXbar)]*1 + [Φ(μ2- βhatXbar) - Φ(μ1-βhatXbar)]*2+[Φ(μ3- βhatXbar) - Φ(μ2-βhatXbar)]*3+ [Φ(μ4- βhatXbar) - Φ(μ3-βhatXbar)]*4 [1-Φ(μ5- βhatXbar)]*5
Φ is the normal CDF (Greene, 2003)
P-values are in parentheses: * indicates significance at 1%; ** at 5% and *** at 10%.

TABLE A2

Dependent Variable= Constructed Health Status (1=excellent to 5=poor)
Ordered Probit Coefficients

Variable	*White*	*Black*	*W/B Pooled*	*W/B Pooled No Smokers*
Constant	0.322	1.023	0.199	0.104
Black			0.095*	0.125*
			(0.00)	(0.00)
Age	0.042*	0.059*	0.045*	0.045*
	(0.00)	(0.00)	(0.00)	(0.00)
Age^2	-0.0004*	-0.0005*	-0.0004*	-0.0004*
	(0.00)	(0.00)	(0.00)	(0.00)
Health Plan	0.068*	0.053**	0.066*	0.044*
	(0.00)	(0.03)	(0.00)	(0.00)
Checkup (never)				
past year	0.119*	0.089*	0.116*	0.104*
	(0.00)	(0.00)	(0.00)	(0.00)
past 2 years	0.036*	-0.029	0.030*	0.027**
	(0.00)	(0.55)	(0.01)	(0.05)
past 5 years	0.040*	0.081	0.042*	0.039*
	(0.00)	(0.20)	(0.01)	(0.01)
past	–0.075*	-0.056	-0.073**	-0.054
	(0.01)	(0.71)	(0.02)	(0.14)
Partner	0.074*	0.050**	0.064*	0.052*
	(0.00)	(0.02)	(0.00)	(0.00)
Education (none)				
1-8^{th}	0.374*	0.103	0.341**	0.355***
	(0.01)	(0.87)	(0.03)	(0.06)
1-11^{th}	-0.112*	-0.107	-0.119*	-0.051
	(0.00)	(0.23)	(0.00)	(0.27)
1-12/GED	-0.099*	-0.036	-0.084*	-0.124*
	(0.00)	(0.36)	(0.00)	(0.00)
1-some college	0.008	0.026	0.015***	0.001
	(0.32)	(0.29)	(0.06)	(0.89)
1-college +	-0.181*	-0.167*	-0.180*	-0.199*
	(0.00)	(0.00)	(0.00)	(0.00)
Income (low)				
10K or more	-0.101*	-0.222*	-0.117*	-0.065**
	(0.00)	(0.00)	(0.00)	(0.03)
15K or more	-0.174*	-0.057	-0.154*	-0.148*
	(0.00)	(0.23)	(0.00)	(0.00)
20K or more	-0.133*	-0.085**	-0.156*	-0.099*
	(0.00)	(0.04)	(0.00)	(0.00)
25K or more	-0.138*	-0.109**	-0.132*	-0.120*
	(0.00)	(0.02)	(0.00)	(0.00)
35K or more	-0.046*	-0.047	-0.046*	-0.058*
	(0.00)	(0.15)	(0.00)	(0.00)
50K or more	–0.062*	-0.018	-0.059*	-0.047*
	(0.00)	(0.59)	(0.00)	(0.00)
High Income	-0.119*	-0.053	-0.119*	-0.133*
	(0.00)	(0.12)	(0.00)	(0.00)

Sex	0.174*	-0.073*	0.149*	0.181*
	(0.00)	(0.00)	(0.00)	(0.00)
Emotional Support (none)				
Rarely	0.265*	0.312*	0.289*	0.308*
	(0.00)	(0.00)	(0.00)	(0.00)
Sometimes	0.083*	0.092**	0.098*	0.129*
	(0.00)	(0.05)	(0.00)	(0.00)
Usually	-0.067*	0.069	-0.032	0.001
	(0.00)	(0.14)	(0.12)	(0.96)
Always	-0.115*	-0.013	-0.080*	-0.040
	(0.00)	(0.76)	(0.00)	(0.11)
Smoker	0.048*	-0.066*	-0.045*	
	(0.00)	(0.01)	(0.00)	
Drinker	0.119*	-0.074	-0.118*	-0.113*
	(0.00)	(0.16)	(0.00)	(0.00)
Fruits/Veg	-0.011*	-0.003	-0.011*	-0.022*
	(0.00)	(0.79)	(0.00)	(0.00)
Exercise (none)				
Insufficient	-0.360*	-0.119*	-0.309*	-0.308*
	(0.00)	(0.00)	(0.00)	(0.00)
Moderate	-0.476*	-0.181*	-0.422*	-0.430*
	(0.00)	(0.00)	(0.00)	(0.00)
Vigorous	-0.583*	-0.269	-0.523*	-0.547*
	(0.00)	(0.00)	(0.00)	(0.00)
Mod+Vig	-0.660*	-0.302*	-0.601*	-0.625*
	(0.00)	(0.00)	(0.00)	(0.00)
High Risk	0.108*	0.128*	0.098*	0.088*
	(0.00)	(0.00)	(0.00)	(0.00)
Midwest	-0.011	0.043	-0.005	0.004
	(0.17)	(0.13)	(0.544)	(0.67)
South	0.021*	0.018	-0.012	-0.011
	(0.01)	(0.40)	(0.13)	(0.22)
West	0.020*	0.052	0.022*	0.016***
	(0.01)	(0.25)	(0.01)	(0.07)
Islands	-0.203*	-0.183*	-0.188*	-0.218*
	(0.00)	(0.00)	(0.00)	(0.00)
Predicted Health Status	2.623016	3.001514	2.649246	2.628331
Sample Size	161,778	16,778	178,556	137,897
Wald Chi2	14089.72	1249.06	14844.61	11913.96
Prob >chi	0.00	0.00	0.00	0.00

Predicted health status is calculated as:
[1-Φ(βhatXbar)]*0 + [Φ(μ1- βhatXbar) - Φ(-βhatXbar)]*1 + [Φ(μ2- βhatXbar) - Φ(μ1-βhatXbar)]*2+[Φ(μ3- βhatXbar) - Φ(μ2-βhatXbar)]*3+ [Φ(μ4- βhatXbar) - Φ(μ3-βhatXbar)]*4 [1-Φ(μ5- βhatXbar)]*5
Φ is the normal CDF (Greene, 2003)
P-values are in parentheses: * indicates significance at 1%; ** at 5% and *** at 10%.