Published: *Journal of Socio-Economics*, 38(6), 2009, pp. 977-87.

Controlling for Endogeneity in the Health-Socioeconomic Status Relationship of the Near Retired

Keith A. Bender* Ioannis Theodossiou**

ABSTRACT

Although there is a strong correlation between health and socio-economic status (SES), the direction of causation is not empirically clear. This study a uses a two stage methodology to control for endogeneity. Using data from the US Health and Retirement Study, the results show that after controlling for endogeneity, labor force status turns out to be an even more important factor in determining health than in the case where endogeneity is not controlled for. In particular, unemployment generates much higher odds for worse health as measured by both subjective and objective health measures.

JEL: I1, J64

Keywords: Socio-economic status, subjective and objective health measures, endogeneity, labor force status

* Corresponding author: Department of Economics and Graduate Program in Human Resources and Labor Relations, PO Box 413, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA. Tel.: 414-229-4761, fax: 414-229-3860, Email address: kabender@uwm.edu.

** Centre for European Labour Market Research and Department of Economics, Aberdeen University, Dunbar Street, Aberdeen, SCOTLAND, UK AB24 3QY. Tel: 44-1224-272183, fax: 44-1224-272181, Email address: theod@abdn.ac.uk.

Acknowledgements: The authors wish to thank seminar participants at the Aberdeen University Centre for European Labour Market Research, the University of Wisconsin-Milwaukee, and the Université Panthéon-Assas Paris II, as well as Dan Hamermesh and Andrew Oswald and an anonymous referee of this journal for helpful comments. Funding for Bender from the UWM Graduate School Research Committee and for Theodossiou from the European Commission, contract number QLK6-CT-2002_02292 is gratefully acknowledged.

Controlling for Endogeneity in the Health-Socioeconomic Status Relationship of the Near Retired

1. Introduction.

Research in European countries and the U.S. has shown an increase in the risks of ill health with decreasing socio-economic status (SES), and there is a growing concern among policy makers that these inequalities in health are widening. Given that many countries (at least partially) fund health care, the link between health and SES is a central one for many policy makers as SES measures like income and wealth have become more unequal over time.

Although medical evidence suggests that the direction of causality runs from SES to health in the sense that lower levels of income lead to worse health outcomes, the causation might also run in the opposite direction. Worse health may mean fewer labor market opportunities because of a weaker attachment to the labor market or directly to lower income because of lower productivity. This endogeneity problem is common in social science research. However, from a policy perspective it is important to disentangle the direction of the relationship since this will provide guidance to policy if increasing a society's well being is a goal for a government.

A further area of policy concern is that of the aging populations in developed countries. In view of the approaching retirement of the 'Baby Boom' demographic cohort, falling fertility rates, and increasing life expectancy, there are strong financial pressures on the public retirement and health systems in developed countries. In the US, because of the public financing of health insurance for older individuals, understanding the relationships between SES and health for older individuals is paramount for policy makers given the above demographic trends.

The aim of this study, therefore, is to address these issues. First, a two stage empirical methodology to control for the endogeneity in the health – SES relationship is presented. In this

paper, two measures of SES, income and labor force status, are used. Second, a nationally representative dataset of older individuals in the US – the Health and Retirement Study (HRS) – is employed. This dataset has a rich set of current and past health and SES measures, in addition to a number of other important control variables. Furthermore, it is important that the dataset provides information on past health and SES outcomes since the literature on the SES-health relationship shows the importance of controlling for past health in explaining current health. The results show that being unemployed or out of the labor force and low income are detrimental to health, even after controlling for differences in individual characteristics and the endogeneity of labor force status and income.

The paper is organized as follows. The first section briefly summarizes the health-SES literature, focusing on the issues relevant for this study. The next section briefly describes the key variables from the HRS dataset. A discussion of the empirical methodology and results follow. A final section concludes.

2. Health and SES.

There is a very large literature, mostly by health professionals and demographers, on the health-SES relationship, and a complete review of this literature is beyond the scope of this paper, since Smith (1999), Auerbach and Krimgold (2001), and Adams *et al*. (2003) offer broad summaries of the literature. Rather, the focus here is on several key issues that are relevant for this study. These issues include a discussion on measures of SES, the health-SES relationship among older individuals, and how the literature has controlled for endogeneity in the relationship.

2.1 Measures of SES.

A number of measures of SES have been proposed in the literature, including income, wealth, labor force status, education, race/ethnicity, and others. The economics literature has been primarily focused on the first three of these. Income and wealth have been found to have important effects on individuals' health since they determine individuals' standards of living. Blaxter (1990) and Ecob and Smith (1999) show that health improves initially as income increases but at the high-income end of the scale, there is a decrease in health. Navarro (1990) further shows that there are great disparities in mortality and morbidity due to disparities in wealth and income even if the effects of race are netted out. Van Doorslaer, *et al*. (1997) find that income inequality is strongly correlated with health inequality across countries. Rahkonen, *et al*. (1997) and Everson, *et al*. (2002) find support for the hypothesis that past and present SES are important determinants of adult health. Lynch, *et al*. (1997) attempt to capture the health effects of prolonged exposure to low income and to evaluate health effects of transitions into and out of low-income groups on different aspects of health. They show that significant associations exist between the number of periods of individual economic adversities and almost all measures of physical and mental functioning. Finally, Adams, *et al*. (2003) control for wealth endogeneity but still find a significant relationship between wealth and certain types of illnesses.

The labor force status of individuals is also found to affect health. In particular, Moser, *et al*. (1986) show that unemployed men and their wives have higher mortality rates, while Blakely, *et al*. (2003) find a link between unemployment and higher suicide rates. Blaxter (1990) finds that they also exhibit lower mental and physical health compared to the employed in each age group, particularly the oldest age group. Ferrie, *et al*. (1995) show that both 'job

change' and 'job insecurity' exhibit negative effects on health status in both genders. Other researchers, e.g. Moser, *et al*. (1984), Dahl (1993), and Bartley (1994), explore aspects of the health-labor force status relationship.

Other research examines the effects of labor force status on psychological health. In an early study, Björklund (1985) finds that cross-section estimates indicate that unemployment leads to worse mental health, although panel data estimates show that controlling for fixed effects generally reduces the statistical significance of unemployment's effects. Clark and Oswald (1994) show that the unemployed suffer almost twice the level of psychological distress of the employed. Theodossiou (1998) finds that unemployment is associated with an increase in anxiety, depression, and loss of confidence, and reductions in self-esteem and the level of general happiness. The study further reveals that the negative psychological consequences of unemployment are, in general, significantly more severe than the psychological consequences of low paid employment or out of the labor force status. More recently, Andersson (2008) finds that self-employment generates increased job satisfaction among workers.

2.2 The Health-SES Relationship at Older Ages.

Most studies of the health-SES relationship do not focus on a particular age group (Lynch, *et al*., 1997, and Van Rossum, *et al*., 2000), although Bender and Habermalz (2008) do examine how the health-SES relationship changes over the life cycle. There are, however, a number of studies that do focus on the relationship for older individuals in society. Evidence regarding differentials in health status across age groups suggests that among older people there is a strong link between relative social standing, educational attainment, occupational status, income, and functional aging. Wilkinson (1986), Blaxter (1990), Grundy (1998), Grundy and

Holt (2000), and Grundy and Gemma (2000) find that the differences in health between classes are very small early in life, but by middle age health differences become large. At later stages of life, class differences in illness and psychological health become pronounced, while for the most objective dimensions of health among the older age groups, such as disease and fitness, converged, although there are some differences across genders. Smith (1999) shows that this relationship still holds even after controlling for the endogeneity of income, and Smith and Kington (1997a) show that the inclusion of controls for specific chronic medical conditions mitigates, but does not eliminate, the impact of SES on functional limitations among older Americans.

Similar patterns hold when investigating mortality. Martelin (1994) focuses on current SES-mortality differentials among the Finnish elderly and shows that there were SES differences in mortality of 4-5 years in the life expectancy at the age of 60 and 2-3 years at the age of 75 between lower and higher SES. In a similar comparison, Van Rossum *et al*. (2000) find that older men in the highest SES have a 5-year survival probability of 93%, while this probability is 84% for their counterparts at the lowest SES.

2.3 The Endogenous Relationship between Health and SES.

Previous research often recognizes the importance of endogeneity. A number of studies control for endogeneity by using panel datasets and information on past health. Thus, Chapman and Hariharan (1994) use a survival model and measures of previous health in regressions of current health to control for causality between wealth and health. Smith and Kington (1997b), use a reduced form model and make a distinction between contemporaneous (current period) feedbacks from health to economic status and health behaviors and the full lifetime sequence of

such feedback relationships. Finally, several other studies, such as Björklund (1985), Gerlach and Stephan (1996), and Winkelmann and Winkelmann (1998) control for endogeneity by controlling for individual fixed effects in panel datasets.

Two studies employ panel models using versions of the HRS data. First, Adams *et al*. (2001) tests Granger causality in the socioeconomic status-mortality relationship. They cannot reject the null hypothesis of no causal link from SES to mortality and to the onset of acute conditions, but they also offer some evidence of a possible link from SES to the gradual onset of chronic conditions. Second, Hurd and Kapteyn (2003) investigate the interrelationship of health, income, and wealth using data from the U.S. Health and Retirement Study and from two Dutch datasets. Based on the cross-section estimations of the effects of income and wealth on health status they calculate the average change in relative risk holding constant wealth by averaging the relative risk over each wealth category. In general, these studies find that a wealth-SES relationship remains even after controlling for endogeneity. It is interesting to note, however, that none of the above studies uses the conventional two stage least square approach to control for the endogeneity of labor force status.

2.4 Summary.

This brief literature review shows several key issues that this paper addresses. First, it focuses on income, labor force status, and wealth as key SES factors. While other SES indicators such as race, ethnicity, and education are important, they are exogenous in this study, and therefore, the focus is on the SES factors which are most likely to be influenced by endogeneity. Second, older members of society are an important demographic group given aging population trends. For this group the SES - health relationship matters most in view of

their approaching vulnerability and frailty in terms of health as age advances. Third, controlling for endogeneity, both in the contemporaneous relationship of health and SES and in the effect of the past health and SES on the current health, is important in understanding the health-SES relationship. The empirical methodology presented below attempts to deal with these issues.

3. Data and Measures of Health Status.

The data are taken from the U.S. Health and Retirement Study (HRS), a nationally representative survey of approximately 10,000 individuals aged 51 to 61 years old in 1992. The HRS contains a range of demographic, income, wealth, and labor force related questions, as well as detailed questions regarding health. In addition, the HRS is a panel dataset, and thus data from the 1996 wave is employed to control for the effects of past health on current (here defined as 1998) health. While the 2000 and 2002 data are available, the study does not employ these data since by 2000 many of the respondents in the HRS are at usual retirement ages. Since, it is the aim of this study to identify the effects of labor market status on health, having a significant proportion of the sample retired would complicate our analysis as it introduces retirement selection problems.[1]

Two sets of health measures are used in this paper. The first set are two self-assessed measures of health.[2] The first question measures overall health and asks, "Would you say your

[1] However, note that early retirement may have occurred in this sample. These individuals are classified as out of the labor force.

[2] Economists are generally suspicious of subjective evaluations, in this case health, issues (see Hamermesh, 2001 for a discussion of this in the job satisfaction literature). However, research by Hurd and McGarry (1997) shows that self-assessed health is highly predictive of mortality. Therefore, it is likely to be an important avenue for exploration of the relationship between aspects of SES and health.

health is excellent, very good, good, fair, or poor?" This variable is coded so that increasing values indicate worse self-assessed health. The second health variable is a measure of emotional health and is based on the following question: "Now think about the past week and the feelings you have experienced. Please tell me if … the following was true for you much of the time during the past week. Much of the time during the past week, you felt depressed. (Would you say yes or no?)" This variable is coded as a binary variable that takes the value of one when the respondent feels depressed and zero otherwise.

The second set of health variables measure actual (but still self reported) health problems of the respondents. The first is a count of the number of serious health illnesses that respondents have. These major illnesses comprise of: currently having high blood pressure, diabetes, cancer, lung disease, and arthritis; ever had a heart attack and/or a stroke; and are on medication for a psychological problem. Two points should be made about this variable. First, since very few have more than four of these major illnesses, this variable has five possible outcomes from zero to four or greater illnesses. Second, a potential problem with counting these illnesses is that it assumes an equal weighting between them. Though this may be unrealistic, it does generate a useful index of ill health (Katz, *et al*., 1963).

The second health outcome measures the presences of a mobility limitation in the respondents. Although it would be desirable to create an index of mobility limitations like Smith and Kingston (1997a), by counting the difficulty that respondents had with at least some difficulty with bathing, dressing, eating, getting into and out of bed, and walking across a room, the responses to this variable are collapsed into a binary variable that indicates whether an individual has some difficulty with any one of these activities because of the infrequency of multiple responses to this count.

Table 1 reports the frequencies of these different measures of the health status for the 1998 data by labor force status and whether income is below the first quartile (approximately $10,000), the middle 50 percent, or above the third quartile (approximately $31,000). The first three columns disaggregate the frequencies by labor force status. For each health measure, those who are working exhibit better measures of health. There are higher frequencies for expressing excellent or very good self assessed physical health, not being depressed, having no or only one illness, and having no mobility restraints. The final three columns split the sample into income categories. They show that as income increases, the percentage of individuals in the better health categories also increases.

(Table 1 around here.)

Clearly, there is some evidence of the correlation between health and labor force status and income but these do not reveal whether they are influenced either by other confounding variables or by reverse causality. In order to control for the former, a number of other potentially interesting determinants of health from the HRS are used. These variables are defined in Appendix Table A1 and their means are reported.

4. Methodology.

As discussed in the literature review above, a common way to deal with endogeneity and to control for past health is through fixed effects models using panel data. While this is a standard method, it suffers from two important limitations First, the method requires that there be sufficient variation in the dependent variable over time. If there is no variation, then observations are dropped from the estimating procedure. Because health changes gradually one should not expect such large variations. In the present dataset, while there is a reasonable

variation of the health outcomes in the ordered response dependent variables (self assessed physical health and the count of illnesses), there is a relatively limited variation for self assessed emotional health (27%) or the presence of a mobility limitation (12%) between the first and fourth waves of the HRS. This would dramatically decrease the sample from which the coefficients are estimated.[3]

A second issue involves the ordered nature of the values of the self assessed health variable. Currently, there is no way to estimate fixed effects models using these types of data. Although one could collapse the dependent variable into a dummy variable (which would allow for a fixed effects estimation via a conditional logit methodology), it would come at the cost of reduced informational value and variability in the dependent variable.

In view of the above limitations, a multistage empirical strategy to correct for endogeneity is employed in this study, akin to the two stage least squares procedure. The methodology is complicated by the fact that two sources of endogeneity are present. First, there is the issue of contemporaneous endogeneity between SES and health. Hence, the use of the standard multivariate regression methodology will only show correlation and not causation. Second, as the review of the literature shows, the influences of past health and SES need to be integrated into the methodology. used to correct for the endogeneity. This section details the way this study controls for both of the above issues.

[3] A related concern is that in a panel framework, one cannot estimate the coefficients on important, but time invariant, independent variables such as gender, race, and education.

4.1 Basic Estimation Issues.

Health is assumed to be a function of income, labor force status, and a set of other covariates with the form of:

$$H_i = (LF)_i \beta_1 + I_i \beta_2 + X_i \beta_3 + \varepsilon_i, \quad (1)$$

where for individual i, H is a measure of health, LF is the labor force status, I is per capita household income (and its square to allow for any nonlinearities), X is a vector of other determinants of health, ε is a random error term of unobservables, β_1, β_2, and β_3 are coefficient vectors to be estimated.

An analysis ignoring endogeneity would estimate eq. (1) to find the estimates of β_1 and β_2. The type of regression methodology depends upon the dependent variable. Since this study utilizes three different types of dependent variable – each requires a different methodology. For the self assessed physical health measure which is ordered, an ordered logistic methodology (see McCullagh, 1980, Anderson and Phillips, 1981, and Ashby *et al*., 1986) is utilized. For the self assessed emotional health and mobility restraint variables, the dependent variable is dichotomous, and hence a standard logit model is used. Finally, the number of illnesses is a count variable, and therefore a Poisson regression methodology is utilized. Since all three methodologies are based on the logistic distribution the estimated coefficients can be converted into log odds ratios which are easily interpretable and offer a good approximation of the marginal effects of a change in an respective independent variable.

4.2 Correcting for Endogeneity.

Endogeneity is an issue because labor force status (LF) and income (I) are also functions of health (H). Linearizing these functions results in the following equations:

$$LF_i = X_i\alpha_1 + Z_i^{LF}\alpha_2 + H_i\alpha_3 + \varepsilon_i^{LF} \quad (2)$$

for the labor force equation and

$$I_i = X_i\gamma_1 + Z_i^{I}\gamma_2 + H_i\gamma_3 + \varepsilon_i^{I} \quad (3)$$

For the income equation, where in addition to the variables specified above, Z^{LF} and Z^{I} are variables that are assumed to affect only labor force status and income but not health. Since H appears in equations (2) and (3), an estimation of eq. (1) will lead to biased estimates.

To control for endogeneity, the approach employed follows the commonly used methodology to correct for endogeneity by using the instrumental variables approach. This involves two stages. First, the equations (2) and (3) are estimated and the predicted values of the LF and I variables are obtained. Second, these predicted values are included in a regression of eq. (1) in place of the actual values of LF and I. Therefore, to control for labor force status endogeneity, a multinomial logit regression is estimated since labor force status could be either working or unemployed or out of the labor force/retired. Using the estimated coefficients from this regression, the probability of being unemployed, working, and out of the labor force/retired are calculated for each respondent. The probabilities of being unemployed and out of the labor force are then included in the health status regressions in the second stage.

A similar procedure is applied in order to generate predictions of an individual's income. In the first stage, a linear regression of per capita household income is estimated. Then, using the estimated coefficients from this regression, income is predicted for each individual. This is then included in the health status regressions.

A final consideration with this approach is to fulfill the identification requirements of the methodology. This entails finding at least one variable (an identifying restriction) to include in the labor force and income regressions that affects labor force and income but not health status,

or the Z^{LF} and Z^{I} vectors from eq.'s (2) and (3). Past values for health, labor force status, and income in the 1996 wave are employed as identifying variables in both the labor force status and income regressions. While these will certainly affect the current values for labor force status and income, it is not possible for the respective current values to affect the past values of these two sets of variables. Furthermore for the labor force status regression, the labor force participation rate and the unemployment rate of the U.S. Census region in which the respondent resides are also used as identifying variables. Although state information on labor force participation and unemployment rates would be preferable, state identifiers do not exist in the public use form of the HRS. For the per capital household income regressions, the number of rooms in the household are used as instruments.[4]

5. Results.

Before turning to the estimates there are two general points to be emphasized. First, for all the estimations, the dependent variables for health are coded so that higher values indicate *worse* health. Therefore, since the coefficients are converted to log odds ratios, coefficients greater than one indicate that an increase in the independent variable is associated with worse health. In the cases of the self assessed emotional health and the presence of a mobility

[4] In the results below, the focus is on four different measures of health. However, in the endogeneity corrections, only the past value of health for the health status variable are included for which the endogeneity effects are of concern for this study. For example, only information on past self assessed health is included when estimating the endogeneity correction for self assessed health and information on past self assessed emotional health is included when estimating the endogeneity correction for current self assessed emotional health. Adding in the all measures of previous health does not change significantly, the corrected estimated results but causes the number of observations to drop since not all HRS respondents answer each health question.

limitation, the interpretation is relatively straightforward since the dependent variables are dichotomous. However, for the other two health measures, the interpretation is somewhat more complicated. Because these dependent variables are ordered in nature, there are several ways of interpreting the log odds ratios depending on which value of the dependent variable one chooses. For the explanation below, we will frame the discussion of as the log odds that a person with a certain characteristic has a change in his or her odds for the highest category (that is, the worse health) compared to not being in that category.

A second general point is that, as much of the health and economics literatures indicate, there are potentially significant differences in health and economic outcomes by gender, and so in addition to the results from the full sample, results disaggregated by gender will also be reported.

5.1 Results for the entire sample.

Table 2 contains the results from the four sets of regressions with no correction for endogeneity in order to establish whether correlations between the health outcomes and the key SES variables exist. However, before turning to this it would be instructive to first examine briefly the correlations between health and the other control variables, most of which are statistically significant and consistent with the literature. After controlling for other variables, males tend to have worse self-assessed health, since the odds are 1.179 times to be in the worst health category compared to women. However the odds of males having poor health are only 0.851 and 0.961 of the females odds for the emotional health and count of illnesses health measures. White respondents exhibit substantially lower odds of poor health compared to those of other races in three of the health measures, except for the number of serious illnesses, where

the odds are slightly higher at 1.072 times. Wealth has a U-shaped relationship with adverse health (except for mobility limitations where there is no statistically significant relationship), although the marginal effect is relatively small given that the wealth variable is in $10,000 units. Overall, this pattern implies that increased wealth lowers the incidence of poor health up to a point when increased wealth leads to worse health. Increased age slightly lowers the odds of worse self assessed physical and emotional health by 0.986 and 0.968 times per year and, unsurprisingly, slightly increases the odds of an increase in the number of serious illnesses by 1.016 times per year. Following the results from the previous literature, marriage is correlated with lower odds of poor health for each measure, sometimes with a large marginal effect such as lowering the odds of being depressed by 0.617 times. Behavioral factors, such as a high value of body mass index (BMI) and smoking history, correspond to higher odds for worse health with relatively large marginal effects such as 1.335 and 1.44 times higher odds for smokers compared to nonsmokers for worse self assessed physical health and the likelihood of mobility limitations. Those with higher levels of education have lower odds of experiencing adverse health, particularly those with a post graduate degree compared to those with no educational degree. People with health insurance have an increase in the odds of poor objective health (as measured by the number of illnesses and the presence of mobility limitations), although since these groups are more likely to need insurance, these results are not likely to be causal but may just reflect a selection process where those with worse health are more likely to demand health insurance.

(Table 2 around here.)

The key variables which are potentially endogenous SES variables – labor force status and income – appear at the top four rows of Table 2. People who are not working face higher odds of worse health for each measure of health. Compared to workers, the unemployed have

double the odds to experience depression and are 2.446 times the odds of the employed to have a mobility limitation, although the relative odds ratios with both self assessed physical health and the number of serious illnesses are not statistically significant. The odds ratio between being out of the labor force and workers for all four measures is statistically significant, even after controlling for all other variables, indicating an increase in the odds of poor health, particularly for the presence of a mobility limitation where the odds are over three times higher. There is also a statistically significant reduction in the odds of poor health as income increases. Like with wealth, as income increases, the odds of poor health decline until income passes a critical point after which the odds of poor health start to increase.

However, the above results are derived without controlling for either the endogenous nature of the health-SES status relationship or the effect of past unemployment or past health status on the current health. Therefore, the methodology outlined above is employed to control for both contemporaneous endogeneity and the influence of past health status and past labor force status. The results from the first stage regressions to control for the contemporaneous endogeneity in labor force status can be found in Appendix Table A2 and in income in Appendix Table A3, although the discussion concentrates on the final stage results which are found in Table 3. Again before turning to the key SES variables, a brief overview of the other covariates is presented. Interestingly, there are some changes compared to the uncorrected results. Except for the self assessed measure of physical health, gender does not change the odds of having worse health. Age is only decreases the odds of poor health for self assessed physical health. BMI and smoking are no longer associated with a change in the odds for depression. In general, education and health insurance have similar correlations to what we found in Table 2.

(Table 3 around here.)

Turning to the effect of the key SES variables, the results show that, in general, the influence of labor force status on current health is not qualitatively different than without the endogeneity correction. In each instance where the unemployment and out of the labor force indicators are statistically significant in Table 2, they are also significant in Table 3 with the addition of coefficient on unemployment for the self assessed physical health measure. Furthermore, for each case the odds for worse health increases for individuals in either labor market status compared to those who are working. The primary difference is that while the odds ratios are higher for being out of the labor force than the ratios for being unemployed in Table 2 for three of the four health measures, when labor force endogeneity is controlled for, the log odds ratios for an increase in the probability of being unemployed are now larger.[5] Therefore, being unemployed corresponds to higher odds of worse health outcomes than either being employed or being out of the labor force, once endogeneity is controlled for. For example, the odds of worse self assessed health are 1.22 times higher for the unemployed than those who are working, 1.19 times higher for depression, and 1.15 times higher for the likelihood of a mobility limitation. This is contrary to the predictions of Ruhm (2000) who finds evidence that unemployment is good for health, but it is in line with the medical literature, such as Stern (1983), Morris *et al*. (1992 and 1994), Hammarström (1994), Morell *et al*. (1993 and 1994), and Wood *et al*. (1999).

The role of per capita household income is somewhat more mixed after the corrections for endogeneity. It is statistically significant only for self assessed physical health and the

[5] However, it should be noted, that the actual value of the log odds ratios cannot be compared across Tables 2 and 3 for the labor force indicators. In Table 2, the labor force variables are dummy variables, while in Table 3, they are predicted probabilities of being unemployed or out of the labor force. Therefore, the marginal impact is an increase in the probability of being unemployed or out of the labor force (compared to an increase in the probability of working).

number of serious illnesses (and only for the squared income squared term), while it is insignificant for emotional health and the presence of mobility limitations. It has also switched signs compared to the previous findings, so that an increase in income increases the odds of having worse health, although the increase is very small for a $1000 increase in income which causes only a 1.007 times increase in the odds. This reflects the usual finding of studies that the social gradient in health within countries is primarily a gradient in social status, rather than a reflection of absolute material living standards (Wilkinson and Pickett, 2006) which in turn is reflected on the evidence that the relationship between Gross National Income per capita and life expectancy not only grows progressively weaker as countries get richer, but disappears altogether among the 25-30 richest (Marmot and Wilkinson, 2001, and Wilkinson, 1997).

5.2 Results disaggregating by gender.

As mentioned above, the health research literature points to health differences across genders. Likewise, the labor economics literature suggests that the factors that determine labor force participation and income are also found to differ by gender. Therefore, one might expect that any relationships between health and SES might also differ by gender. To allow for this possibility, the health regressions were estimated by gender thus to allow the estimated coefficients to differ. Table 4 reports the results from the disaggregated regressions corrected for endogeneity (results from uncorrected estimates are available from the authors). The main conclusion from these results is that pooling the genders in the same sample masks important quantitative, although not qualitative, differences in the health-labor force status relationship. Qualitatively, the effect of an increase in the probability of being unemployed or out of the labor force (compared to being employed) increases the odds of poor health in nearly every case (the

exception of the probability of being in unemployment for the number of illnesses health measure for both genders and for females in the presence of a mobility limitation). The increase in odds for worse health are relatively small when comparing those out of the labor force compared to workers for both genders, with the increase in odds ranging from 1.004 times (in the number of serious illnesses regression for women) to 1.018 times (for self assessed physical health for men). Presumably, this small effect is due to the choices of not participating in the labor force if one has poor health.

On the other hand, the increase in odds of worse health are relatively large for the unemployed, particularly for men since the odds are higher for unemployed men compared to unemployed women, in each case where the unemployment coefficient is statistically significant. For example, the odds of an unemployed male having worse self assessed health is 1.289 times the odds for a working male, while for unemployed females, the odds increase only 1.174 times. The difference is even greater for the likelihood for depression where the increase in odds is 1.341 times for males and 1.135 times for females. This implies that unemployment is generally worse for men than for women with respect to health, in line with the literature, such as Ferrie *et al*. (1995) and Theodossiou (1998).

As with the full sample results, the effects of income on health by gender are somewhat mixed. Income plays a statistically significant role for females in affecting the self assessed physical health, the number of serious illnesses, and the presence of mobility limitations (although only at the 10 percent level for the latter for the squared term). However, predicted income is has an opposite effect on the first two of these illnesses with increases in income having an inverted U-shape for self assessed physical health but a U-shaped relationship for the number of serious illnesses. In general, income seems to play little role for males in the

endogeneity corrected results. The only exception is self assessed physical health, where we find the same inverted U-shaped relationship, although the effect is about a third of the size as for females.

6. Conclusions

Using a multistage estimation technique to statistically control for endogeneity of labor force status and income, this paper examines the effect of a number of key variables in determining various facets of both self-assessed and physical health. Even after controlling for endogeneity, the results are consistent with the health and economics literature. Overall, the results suggest that unemployment and being out of the labor force have significantly detrimental effects on health status. The results with respect to income show that changes in income are not always significantly related to health. The results further show that the effects of unemployment and being out of the labor force have larger effects on the health status for men than women.

As a sizable proportion of the population of developed countries begins to retire, issues of health, SES, and income will be central to individuals and policy makers alike. While these issues have been the focus of much research, many studies do not address the fact that the endogeneity may cause biases in the estimated size of the correlations between them to the extent that it substantially overestimates gender differences on the effects of SES and income on the different facets of the health status of the older labor force. Therefore, correctly identifying and controlling for the simultaneous nature of the health, SES, and income relationship is central for understanding the importance of the relationship and for guiding policy.

References

Adams, P., Hurd, M. D., McFadden, D., Merrill, A., Ribeiro, T., 2003. Healthy, wealthy and wise? Tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics*, 112, 3-56.

Anderson, J. A., Phillips, P. R., 1981. Regression discrimination and measurement models for ordered categorical variables. *Applied Statistics*, 30, 22-31.

Andersson, P., 2008. Happiness and health: Well-being among the self-employed. *Journal of Socio-Economics*, 37, 213-236.

Ashby, D., Pocock, J., Shaper, A.G., 1986. Ordered polytomous regression: An example relating serum biochemistry and hematology to alcohol consumption. *Applied Statistics*, 35, 289-301.

Auerbach, J. A., Krimgold, B. K., eds., 2001. *Income, Socioeconomic Status and Health: Exploring the Relationships*, Washington, DC: National Policy Association.

Bartley, M., 1994. Unemployment and ill health: Understanding the relationship. *Journal of Epidemiology and Community Health*, 48, 333-337.

Bender, K.A., Habermalz, S., 2008. Are there differences in the health-socioeconomic status relationship over the life cycle? Evidence from Germany. *LABOUR*, 22, 107-25.

Björklund, A., 1985. Unemployment and mental health: Some evidence from panel data. *Journal of Human Resources*, 20, 469-483.

Blakely, T.A., Collings, S.C., Atkinson, J., 2003. Unemployment and suicide. Evidence for a causal association? *Journal of Epidemiological and Community Health*, 57, 594-600.

Blaxter, M., 1990. *Health and Lifestyles*. London: Tavistock/Routledge.

Chapman K. S., Hariharan, G., 1994. Controlling for causality in the link from income to mortality. *Journal of Risk and Uncertainty*, 8, 85-93.

Clark, A. E., Oswald, A.J., 1994. Unhappiness and unemployment. *Economic Journal*, 104, 648-659.

Dahl, E., 1993. Social inequality in health – the role of the healthy worker effect. *Social Science & Medicine,* 36, 1077-1086.

Ecob, R., Davey Smith, G., 1999. Income and health: What is the nature of the relationship? *Social Science & Medicine,* 48, 693-705.

Everson, S.A., Maty, S.C., Lynch, J.W., Kaplan, G.A., 2002. Epidemiologic evidence for the relation between socioeconomic status and depression, obesity, and diabetes. *Journal of Psychosomatic Research,* 53, 891-5.

Ferrie, J.E., Shipley, M.J., Marmot, M.G., Stansfeld, S., Smith, G.D., 1995. Health effects of anticipation of job change and non-employment: Longitudinal data from the Whitehall II Study. *British Medical Journal*, 311, 1264-1269.

Gerlach, K., Stephan, G., 1996. A paper on unhappiness and unemployment in Germany. *Economic Letters*, 52, 325-330.

Grundy, E., 1998. Ageing, ill health and disability. In Tallis, R. (Ed.), *Increasing Longevity: Medical, Social and Political Implications*. London: Royal College of Physicians of London., pp. 16-30.

Grundy, E., Holt, G., 2000. Adult life experiences and health in early old age in Great Britain. *Social Science & Medicine*, 51, 1061-1074.

Grundy, E., Gemma, H., 2000. Adult life experiences and health in early old age in Great Britain. *Social Science & Medicine,* 51, 1061-1074.

Hamermesh, D. S., 2001. The changing distribution of job satisfaction. *Journal of Human Resources*, 36, 1-30.

Hammarström, A., 1994. Health consequences of youth unemployment - review from a gender perspective. *Social Science & Medicine*, 38, 699-709.

Hurd, M., Kapteyn, A., 2003. Health, wealth, and the role of institutions. *Journal of Human Resources*, 38, 387-415.

Hurd, M., McGarry, K., 1997. Evaluation of the subjective probabilities of survival in the Health and Retirement Study. *Journal of Human Resources*, 30, S268-292.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., Jaffe, M.W., 1963. Studies of illness in the aged: the index of ADL – a standard measure of biological and psychosocial function. *Journal of the American Medical Association*, 185, 914-919.

Lynch, J. W., Kaplan, G. A., Shema, S. J., 1997. Cumulative impact of sustained economic hardship on physical, cognitive, psychological, and social functioning. *New England Journal of Medicine*, 337, 1889-1895.

Marmot, M., Wilkinson, R. G., 2001. Psychosocial and material pathways in the relation between income and health. *British Medical Journal*, 322, 1233–1236.

Martelin, T., 1994. Mortality by indicators of socioeconomic status among the Finnish elderly. *Social Science & Medicine*, 38, 1257-1278.

McCullagh, P., 1980. Regression models for ordinal data. *Journal of the Royal Statistical Society, Series (B)*, 42,109-142.

Morris, J. K., Cook, D. G., Shaper, A. G., 1994. Loss of employment and mortality. *British Medical Journal*, 308, 1135-1139.

Morris, J. K., Cook, D. G., Shaper, A. G., 1992. Non-employment and changes in smoking, drinking, and body weight. *British Medical Journal*, 304(6826), 536-541.

Morrell, S., Taylor, R., Quine, S., Kerr, C., 1993. Suicide and unemployment in Australia, 1907-1990. *Social Science & Medicine*, 36(6), 749-756.

Morrell, S., Taylor, R., Quine, S., Kerr, C., Western, J., 1994. A cohort study of unemployment as a cause of psychological disturbance in Australian youth. *Social Science & Medicine*, 38(11), 1553-1564.

Moser, K.A., Fox, J., Jones, D. R., 1984. Unemployment and mortality: further evidence from the OPCS longitudinal study. *The Lancet*, 365-367.

Moser, K.A., Fox, J., Jones, D. R., 1986. Unemployment and mortality in the OPCS longitudinal study. In Wilkinson, R.G. (Ed.). *Class and Health: Research and Longitudinal Data*, London: Tavistock.

Navarro, V., 1990. Race or class versus race and class: mortality differentials in the United States. *The Lancet*, 1238-1240.

Rahkonen, O., Lahelma, E., Huuhka, M., 1997. Past Or present? Childhood living conditions and current socio-economic status as determinants of adult health. *Social Science & Medicine*, 44, 327-36.

Smith, J. P., 1999. Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives*, 13, 145-166.

Smith, J. P., Kington, R. S., 1997a. Demographic and economic correlates of health in old age. *Demography*, 34, 159-170.

Smith, J. P., Kington, R. S., 1997b. Race, socioeconomic status and health in late life. In Martin, L.G., Soldo, B.J. (Eds.) *Racial and Ethnic Differences in the Health of Older* Americans, Washington, D.C.: National Academy Press, pp. 106-162.

Stern, J., 1983. The relationship between unemployment, morbidity and mortality in Britain. *Population Studies*, 37, 61-74.

Theodossiou, I., 1998,. The effects of low-pay and unemployment on psychological well-being: a logistic approach. *Journal of Health Economics*, 17, 85-104.

Van Doorslaer, E., Wagstaff, A., Bleichrodt, H., Calonge, S., Gerdtham, U. G., Gerfin, M., Geurts, J., Gross, L., Hakkinen, U., Leu, R. E., O'Donnell, O., Propper, C., Puffer, F., Rodriguez, M., Sundberg, M., Winkelhake, O., 1997. Income-related inequalities in health: some international comparisons. *Journal of Health Economics*, 16, 93-112.

Van Rossum, C.T.M., Van de Mheen, H., Mackenbach, J.P., Grobbee, D.E., 2000. Socioeconomic status and mortality in Dutch elderly people-The Rotterdam study. *European Journal of Public Health*, 10, 255-261.

Wilkinson, R. G., 1986. Income and mortality. In Wilkinson, R.G. (Ed.) *Class and Health: Research and Longitudinal Data*, London: Tavistock.

Wilkinson, R. G., 1997. Income, inequality and social cohesion. *American Journal of Public Health*, 87: 104–106.

Wilkinson, R. G., Pickett, K. E., 2006. Income inequality and health: A review and explanation of the evidence. *Social Science & Medicine*, 62, 1768-84.

Winkelmann, L., Winkelmann, R., 1998. Why are the unemployed so unhappy? Evidence from panel data. *Economica*, 65, 1-15.

Wood, E., Sallar, A. M., Schechter, T., Hogg, R.S., 1999. Universal health care? Center for Health Evaluation and Outcome Sciences, Monograph Series No. 1. UK.

Table 1.
Frequencies of the Health Status Measures by Labor Force Status and Income

	Labor Force Status			Income		
Health Measure	Working	Unemployed	Out of the labor force	Bottom Quartile	Middle 50%	Top Quartile
Self Assessed Physical Health						
Excellent	17.0%	14.4%	10.6%	89.2%	12.0%	22.1%
Very good	34.1	24.4	25.8	20.4	31.3	37.3
Good	33.7	32.3	31.3	30.5	35.3	30.0
Fair	13.0	22.6	20.5	25.5	16.5	8.4
Poor	2.2	6.3	11.7	14.3	4.9	2.1
Self Assessed Emotional Health						
Not Depressed	88.1	72.9	81.8	74.2	86.5	92.1
Depressed	11.9	27.1	18.2	25.8	13.5	7.9
Count of Illnesses						
0	27.8	31.5	16.3	19.8	20.8	28.7
1	35.6	27.7	29.8	26.5	33.9	37.0
2	23.2	19.6	25.9	24.4	25.2	22.8
3	9.8	6.9	16.2	15.2	13.5	8.7
4+	3.6	14.4	11.9	14.1	6.6	2.8
Mobility Restraint?						
No	95.5	86.9	85.2	82.4	92.0	96.2
Yes	4.5	13.1	14.8	17.6	8.0	3.8

Notes: Data are from the 1998 wave of the HRS. Frequencies are weighted by HRS sample weights.

Table 2.
HRS 1998 Wave Ordered Logistic Regressions (No Endogeneity Correction)

	Measure of Health			
Variables	Self Assessed Physical Health	Self Assessed Emotional Health (Depressed)	Number of Serious Illnesses	Presence of Mobility Limitation
Unemployed	1.361	2.038**	1.094	2.446**
Not in LF	1.910***	1.470***	1.279***	3.070***
Household income	0.899***	0.916***	0.980***	0.821***
Income squared	1.003**	1.002***	1.000***	1.003***
Male	1.179***	0.851*	0.961**	1.002
White	0.634***	0.678***	1.072**	0.720**
Black	0.849	0.918	1.088**	0.999
Household wealth per capita	0.960***	0.918**	0.984***	0.949
Wealth Squared	1.000***	1.000**	1.000***	1.000
Age	0.986*	0.968**	1.016***	0.976
Married	0.789***	0.617***	0.913***	0.751***
Body Mass Index	1.060***	1.002	1.031***	1.055***
Ever smoked	1.335***	1.161*	1.130***	1.440***
High school diploma	0.478***	0.572***	0.890***	0.731***
Bachelor degree	0.351***	0.510***	0.822***	0.444***
Post graduate degree	0.302***	0.350***	0.823***	0.463***
Any Insurance	1.071	0.996	1.107***	1.438***
Log likelihood	-9967.7	-2594.5	-28645981	-1934.4
Chi2(25)	2290.0***	361.7***	1228.7***	410.4***
Number of observations	7211	6678	7216	7192

Note: *, **, *** indicate a significant improvement in the log-likelihood at 10, 5 and 1 percent level, respectively. Cut points or constant terms, eight dummy variables for region, and standard errors are also estimated but not reported. Log odds ratios are reported for each regression.

Table 3.
HRS 1998 Wave Ordered Logistic Regressions Controlling for Past Health and SES. (Endogeneity Correction)

	Measure of Health			
Variable	Self Assessed Physical Health	Self Assessed Emotional Health (Depressed)	Number of Serious Illnesses	Presence of Mobility Limitation
Probability unemployed	1.220***	1.194***	0.990	1.152**
Probability not in LF	1.015***	1.008***	1.006***	1.016***
Predicted household income	1.007**	1.077	0.976	0.755
Predicted income squared	1.000**	0.999	1.001***	1.004
Male	1.373***	0.904	1.018	1.141
White	0.698***	0.721**	1.032	0.840
Black	1.039	1.127	1.049	1.122
Household wealth per capita	0.951***	0.893***	0.979***	0.950
Wealth Squared	1.000***	1.000***	1.000***	1.000
Age	0.980**	0.976	1.001	0.973
Married	0.833***	0.658***	0.897***	0.815*
Body Mass Index	1.060***	1.003	1.030***	1.058***
Ever smoked	1.272***	1.088	1.131***	1.368***
High school diploma	0.492***	0.553***	0.929***	0.796
Bachelor degree	0.333***	0.421***	0.892**	0.491**
Post graduate degree	0.305***	0.288***	0.874**	0.593
Any Insurance	1.155**	1.152	1.081**	1.680***
Log likelihood	-9548.0	-2487.7	-27399956	-1882.0
Chi2(25)	1215.7***	344.9***	1536.0***	365.1***

Note: *, **, *** indicate a significant improvement in the log-likelihood at 10, 5 and 1 percent level, respectively. Cut points, eight dummy variables for region, and standard errors are also estimated but not reported. Log odds ratios are reported for each regression.

Table 4.
HRS 1998 Wave Ordered Logistic Regressions Controlling for Past Health and SES: Differences by Gender. (Endogeneity correction)

	Measure of Health							
	Self Assessed Physical Health		Self Assessed Emotional Health (Depressed)		Number of Serious Illnesses		Presence of Mobility Limitation	
Variable	Females	Males	Females	Males	Females	Males	Females	Males
Probability unemployed	1.174***	1.289***	1.135**	1.341***	0.994	0.957	1.085	1.257**
Probability not in LF	1.014***	1.018***	1.007***	1.009***	1.004***	1.006***	1.016***	1.015***
Predicted household income	1.026***	1.008**	1.043	1.123	0.875***	0.976	0.585	0.744
Predicted income squared	1.000***	1.000**	1.000	0.998	1.005***	1.001*	1.007*	1.003
White	0.549***	0.911	0.666**	0.816	1.096*	1.012	0.946	0.902
Black	0.917	1.148	1.082	1.171	1.070	1.020	1.133	1.013
Household wealth per capita	0.941***	0.949**	0.889**	0.893**	0.988	0.982*	0.783***	1.099*
Wealth Squared	1.000***	1.000**	1.000**	1.000**	1.000	1.000*	1.001***	1.000
Age	0.979*	0.978	0.975	0.985	1.005	0.995	0.952**	1.006
Married	0.778***	0.945	0.712***	0.593***	0.935**	0.913**	0.780	1.118
Body Mass Index	1.063***	1.048***	1.003	1.008	1.029***	1.030***	1.067***	1.021
Ever smoked	1.294***	1.226**	1.179	0.944	1.143***	1.120***	1.421***	1.252
High school diploma	0.455***	0.549***	0.588***	0.491***	0.919**	1.005	0.866	0.907
Bachelor degree	0.262***	0.412***	0.497***	0.338***	0.948	0.940	1.130	0.254***
Post graduate degree	0.259***	0.344***	0.345***	0.230***	0.954	0.924	0.592	0.710
Any Insurance	1.196**	1.073	1.147	1.174	1.145***	1.058	2.017***	1.490
Log likelihood	-5097.2	-4419.1	-1534.9	-949.5	-14619669	-12727699	-1025.2	-815.2
Chi2(24)	965.6***	500.2***	190.8***	152.0***	1091.9***	602.8***	312.2***	164.2***

Note: Numbers are log odds ratios of coefficients. *, **, *** indicate a significant improvement in the log-likelihood at 10, 5 and 1 percent level, respectively. Cut points, eight dummy variables for region, and standard errors are also estimated but not reported.

Appendix Table A1.
Definitions and Means of the Variables

Variable	Definition	Average
Self Assessed Physical Health	Self assessment of physical health status	2.707
Self Assessed Emotional Health	Self assessment of whether respondent is depressed or not	0.150
Number of Serious Illnesses	Number of serious illnesses	1.485
Presence of Mobility Limitations	Presence of mobility limitation	0.092
Unemployed	Equal to 1 if the respondent is unemployed.	0.011
Not in LF	Equal to 1 if the respondent is not in the active labor force.	0.437
Employed	Equal to 1 if the respondent is employed (excluded in the regressions)	0.552
Household income per capita	Divided by 1000.	29.378
Male	Equal to 1 if the respondent is a male.	0.473
White	Equal to 1 if the respondent is white.	0.824
Black	Equal to 1 if the respondent is black.	0.087
Other races	Equal to 1 if the respondent is neither back nor white (excluded in the regressions)	0.021
Household wealth per capita	Divided by 10,000.	19.567
Age	Age of the respondent in years.	61.278
Body Mass Index	Respondent's weight in kilograms/(height in meters)^2	27.350
Ever smoked	Equal to 1 if the respondent ever smoked.	0.629
No qualifications	Equal to 1 if the respondent has no qualifications (excluded in the regressions).	0.210
High school diploma	Equal to 1 if highest degree is a High school diploma.	0.550
Bachelor degree	Equal to 1 if highest degree is a bachelor degree.	0.113
Post graduate studies	Equal to 1 if highest degree is a post graduate degree.	0.088
Married	Equal to 1 if the respondent is married.	0.713
Any Insurance	Equal to 1 if the respondent has health insurance.	0.821

Appendix Table A2.
Multinomial Logit Results of Labor Force Participation

	Self Assessed Physical Health		Self Assessed Emotional Health (Depressed)		Number of Serious Illnesses		Presence of Mobility Limitation	
	Unempl.	Out of LF	Unempl.	Out of LF	Unempl.	Out of LF	Unempl.	Out of LF
Male	-0.232	-0.509***	-0.239	-0.488***	-0.230	-0.430***	-0.237	-0.468***
White	-0.529	0.272**	-0.373	0.113	-0.608*	0.071	-0.573	0.141
Black	-0.777*	0.191	-0.722	0.097	-0.828*	0.046	-0.816*	0.108
Wealth	-0.180	0.057**	-0.173	0.052**	-0.179	0.051**	-0.181	0.045*
Wealth squared	3.9E-04	-0.002***	3.7E-04	-0.002**	3.9E-04	-0.002**	3.9E-04	-0.002**
Age	-0.062	0.161***	-0.052	0.156***	-0.061	0.148***	-0.064	0.153***
BMI	-0.004	0.010	0.001	0.018**	-0.001	0.005	-0.003	0.012
Ever smoked	0.147	0.100	0.238	0.127**	0.148	0.088	0.160	0.120
HS diploma	0.002	-0.318***	-0.124	-0.498***	-0.055	-0.463***	-0.036	-0.464***
Bachelor degree	0.320	-0.450***	0.223	-0.686***	0.226	-0.642***	0.277	-0.647***
Postgrad degree	-0.364	-0.230	-0.434	-0.482***	-0.432	-0.440***	-0.409	-0.443***
Any insurance	-0.387	0.322***	-0.485*	0.306***	-0.430	0.295***	-0.428	0.320***
Married	-0.490**	0.258***	-0.430	0.239***	-0.480*	0.248***	-0.497**	0.219**
Unemp rate in region	-0.071	0.025	-0.032	0.035	-0.063	0.040	-0.063	0.035
LFP rate in region	-0.303	-0.585	-0.024	0.005	-0.031	-0.004	-0.029	-0.009
Unemployed in 1996	1.340***	1.192***	1.307**	1.079***	1.382***	1.120***	1.347***	1.116***
Out of LF in 1996	1.200***	3.338***	1.163***	3.362***	1.216***	3.336***	1.245***	3.386***
Income in 1996	-0.060	-0.260*	-0.142	-0.304*	-0.092	-0.331**	-0.081	-0.307*
V. good health in 1996	0.264	0.172						
Good health in 1996	0.006	0.266**						
Fair health in 1996	0.740*	0.993***						
Poor health in 1996	0.732	2.552***						
Depressed in 1996			0.559*	0.527***				
One illness in 1996					-0.314	0.200**		
Two illnesses in 1996					-0.210	0.463***		
Three illnesses in 1996					-0.183	0.688***		

Four+ illnesses in 1996	0.732	1.426***		
Mobility restraint in 1996			0.647	1.222***

Note: *, **, *** indicate a significant improvement in the log-likelihood at 10, 5 and 1 percent level, respectively. Constant terms, regional indicators, and standard errors are also estimated but not reported. Coefficients are in relation to being employed. Reference variables for grouped independent variables are (where applicable): other race, no educational degree, working in 1996, excellent health in 1996, and no illnesses in 1996.

Appendix Table A3.
OLS Regression of Per Capital Household Income

	Self Assessed Physical Health	Self Assessed Emotional Health (Depressed)	Number of Serious Illnesses	Presence of Mobility Limitation
Constant	8.573***	8.567***	8.660***	8.467***
Male	0.032	0.027	0.026	0.024
White	0.464***	0.461***	0.556***	0.488***
Black	0.141*	0.124	0.178**	0.163**
Wealth	0.059***	0.060***	0.057***	0.060***
Wealth squared	-1.4E-4***	-1.4E-4***	-1.3E-4***	-1.4E-4***
Age	-0.001	-0.001	-0.001	-1.3E-4
BMI	-0.003	-0.005	-0.005*	-0.004
Ever smoked	0.038	0.034	0.042	0.031
HS diploma	0.397***	0.438***	0.474***	0.425***
Bachelor degree	0.492***	0.534***	0.693***	0.528***
Postgrad degree	0.696***	0.739***	0.907***	0.735***
Any insurance	0.392***	0.382***	0.395***	0.392***
Married	0.267***	0.256***	0.294***	0.273***
Rooms=1	-0.026	-0.042	-0.117**	-0.037
Rooms=2	-0.134	-0.128	-0.247**	-0.143
Rooms=3	-0.073	-0.082*	-0.190***	-0.081*
Rooms=4	-0.014	-0.024	-0.104**	-0.022
Unemployed in 1996	-0.682***	-0.667***	-0.726***	-0.681***
Out of LF in 1996	-0.382***	-0.409***	-0.432***	-0.406***
Income in 1996	0.907***	0.912***	0.278	0.912***
V. good health in 1996	-0.003			
Good health in 1996	-0.042			
Fair health in 1996	-0.234***			
Poor health in 1996	-0.374***			
Depressed in 1996		-0.187***		
One illness in 1996			0.063*	
Two illnesses in 1996			0.001	
Three illnesses in 1996			-0.086*	
Four+ illnesses in 1996			-0.303***	
Mobility restraint in 1996				-0.252***

Note: *, **, *** indicate a statistical significance at the 10, 5 and 1 percent level, respectively. Constant terms, regional indicators, and standard errors are also estimated but not reported. Reference variables for grouped independent variables are (where applicable): other race, no educational degree, number of rooms>4, working in 1996, excellent health in 1996, and no illnesses in 1996.