Frijters, Paul and Ulker, Aydogan (2007) The Robustness of the Different Health Measures with Respect to Life Style Choices. In *Proceedings 2007 Australasian Meeting of the Econometric Society*, Brisbane.

THE ROBUSTNESS OF THE DIFFERENT HEALTH MEASURES WITH RESPECT TO LIFE STYLE CHOICES

Paul Frijters[*] and Aydogan Ulker[**]

** School of Economics, Queensland University of Technology, Brisbane, Australia*
[**] RSSS, Australian National University, Canberra, Australia

January 2007

Abstract

Do different health variables measure the same thing? In this paper we investigate the robustness of the effects of life-style choices on (1) self-assessed general health status, (2) problems with undertaking daily tasks and chores, (3) mental health indicators, (4) BMI, (5) the presence of serious long-term health conditions, and (6) mortality. The lifestyle choices we consider are regular exercise, being a smoker and the amount of alcohol consumed. We furthermore distinguish between short-run effects and long-run effects, and estimate both ordinal models and cardinal models. We estimate the models using longitudinal data drawn from the US Health and Retirement Study (HRS) between 1992 and 2002. We find surprisingly large differences in effects of lifestyle on the health measures and a general lack of consistency between our measures. Exercise is found to significantly reduce mortality both in the short and long-run, but it has little effect on stated health or doctor-assessed health measures. Importantly, smoking is found to have a long-run effect on mortality, but smoking improves self-stated health, reduces the problems individuals have with doing daily chores, improves mental health, reduces the number of measured serious illnesses and reduces Body Mass Index (BMI). Finally, we find no short-run or long-run benefits of income or wealth, implying that the effect of wealth and income would have to work via the increased levels of exercise and reduced levels of smoking associated with higher income and wealth levels.

Keywords: Morbidity, Mortality, Lifestyle, Alcohol, Smoking, Exercise, Income, Wealth

JEL Classifications: Z1, C23, C25, I31

Corresponding author's name: Aydogan Ulker, address: Economics Program, Research School of Social Sciences, Bldg. 9, Australian National University, Canberra ACT 0200, Australia, e-mail: aulker@coombs.anu.edu.au, phone: +61 (0)2 6125 3198, fax: +61 (0)2 6125 0182.

1. Introduction

While hundreds of studies have examined the relationship between health outcomes and socio-economic factors such as income and lifestyle choices[1], there remains considerable uncertainty about the short and long-term determinants of health. Take the case of smoking and drinking. The consensus in the literature is that cigarette smoking is the leading cause of lung cancer deaths, the leading cause of chronic bronchitis and a significant cause of heart disease and stroke (U.S. Department of Health and Human Services (1989)). Ezzati and Lopez (2003) estimate from the incidence of lung-cancer that there were up to 5 million premature smoking-related deaths in 2000 in the world. Many of the hazards of smoking are long-term in the sense that smoking has been found to have a cumulative effect (Peto (1986)). There's much less consensus on whether smoking has a strong short-run effect or whether smoking affects self-reported health measures in the same way as mortality. For drinking, there is no consensus about any effect. Di Castelnuovo et. al. (2006) perform a meta-analysis of prospective studies on alcohol dosing and total mortality and find that low levels of alcohol intake (1-2 drinks per day for women and 2-4 drinks per day for men) are inversely associated with mortality in both men and women. Moderate drinking (2 drinks per day for men and 1 drink for women) is associated with relatively low mortality rates in the US (Liao et. al. (2000)) and the Netherlands (San Jose et. al. (2000)), and particular health benefits from drinking are reported for coronary disease and type 2 diabetes (Caruso et. al. (2000)). Yet, this association does not necessarily imply causality: Knupfer (1987) finds that drinking very infrequently is just as good in health terms as drinking moderately, implying that the found benefits from drinking in other studies may be due to reverse causality, i.e. people only keep drinking if they are still healthy. Indeed, Sturm (2002), using data from the US Health Care for Communities survey, finds significant negative influences from average drinking on physical health scales and on a number of chronic health conditions. These studies mainly relate to short-term effects though, implying that the long-run effects from average drinking are still elusive.

A key issue in the economic health literature is whether different health measures tell the same story. For example, if self-stated health and mortality are affected by the same life style variables to the same degree, then one can use results on either one of them to draw

[1] Recent contributions looking at the relation between health and a variety of lifestyles and financial measures include Adams et al. (2003); Attanasio and Hoynes (2000); Benzeval et. al. (2000); Benzeval and Judge (2002); Burstrom and Fredlund (2001); Case at al. (2004); Gerdtham and Johannesson (2004); Meer et al. (2003); Ruhm, (2000); Jones et al. (2004).

inferences about the other. If, on the other hand, different health measures tell very different stories then one has to be far more careful in drawing general inferences about 'health' from looking at any single or even at a group of health variables.

Since most empirical researchers do not have input into the construction of the datasets they use, the health measures that they use are often not the ones that they might have chosen in an ideal world. Typically, cross-section and panel datasets will have only a few health questions, and these questions will differ from survey to survey. Whether different health measures tell the same story is then important for ascertaining the limits of what one can learn about health in general from such surveys. Cross-sectional surveys for instance invariably have no information on those that die and hence inferences on mortality from cross-sectional data can only be made if there are health measures that are not just predictive of mortality but that are affected in roughly the same way by every health-related choice. Panel surveys differ amongst each other in their coverage of health issues and differ in the wording of their health questions. For example, the US PSID panel dataset and the German Socio Economic Panel have self-stated health every year but not doctor-assessed health or any of the other measures we look at in this paper. The European Community Household Panel (ECHP) includes measures of mental health and self-reported physical limitations, but does not yet appear to have good mortality data. The British Household Panel Survey (BHPS) includes both self-assessed health, a measure of mental health, a list of physical disabilities, and a general health question based on 12 items ranging from sleep deprivation to being stressed (the GHQ 12). The BHPS, however, lacks doctor-assessed health. In a strict sense, the questions posed differ between these panels. The self-assessed health variable in the BHPS for instance asks people to rate themselves relative to others of the same age whereas the self-assessed health variable in the ECHP is not age-related. Furthermore, wordings change over time within the same survey, such as for the BHPS where self-assessed health in the 9th wave was not related to 'health in the last 12 months' whereas self-assessed health in all other waves was related to the last year. Since these panels are amongst the mostly widely used ones in economic analyses of health, the generalisability of findings on any of the measures in one of these datasets matters for the inferences we can draw.

In this paper we assess the robustness of the relationship between on the one hand life style choices and material indicators and on the other hand, various measures of health from the same dataset. We use longitudinal data drawn from the US Health and Retirement Study (HRS) between 1992 and 2002 which contains a large number of different health measures as well as good mortality data and is thus ideal for the question at hand. We will use 6 different health measures: stated health (ranging from excellent to very bad); mental health (which is made up of 8 questions concerning mental wellbeing); difficulties performing several chores (11 items such as walking around and clothing); Body Mass Index (BMI); doctor-assessed serious illnesses (8 items, including whether a person has had cancer and diabetes); and, finally, mortality. We will interpret mortality to be the most reliable arbiter of whether someone's health was good or bad. The lifestyle choices we consider are regular exercise, being a smoker and the amount of alcohol consumed.

The advantage of using a single data source to assess robustness is that we don't have to worry that our results are biased due to comparing different types of individuals or due to differing general socio-economic circumstances. It also means that one can use the same control variables for each individual health outcome. This matters because many variables are correlated with each other implying that studies can get different findings depending on their set of control variables. It is for instance the case that the wealthier on average engage more in exercise and are less likely to be obese (see Cutler and Glaeser 2005). If one would study wealth and health without taking obesity and exercise into account, one would find an inflated effect of wealth that one could only interpret as a valid causal finding if one would argue that higher wealth causes more exercise and lower obesity. In terms of robustness, one would find a different 'wealth effect' across two different data sets for two different health measures if one controlled for exercise in one data set but not in the other, even if the two health measures would be affected by the same variables in exactly the same way.

In Section 2 we will briefly review the health studies that have used the same dataset as this study in order to establish some background to our robustness exercise. In Section 3 we detail the data used and pay specific attention to our health measures.

In Section 4 and 5 we look at robustness across two dimensions. In Section 4 we look at whether the sign and statistical significance of the found coefficients of lifestyle choices for various health measures correspond with each other. To this end we use fixed effects ordinal models that take out any persistent individual traits, and we simply compare the coefficients of the various life style choices across the health measures. There, we find that the negative effect of smoking is only picked up by mortality and not by any other of these health measures. It turns out that not only does smoking positively affect the non-mortality measures in the short-run (which could be explained by selection effects) but that even the long-run effect of smoking on the non-mortality measures is not negative. The effects of income and wealth are robust across measures in the sense that they have very small and usually insignificant effects for the health measures.

In Section 5 we look at robustness from the point of view of how well the health measures actually capture the factors that are important for mortality and how much of the relation between various life style choices and mortality runs through those health measures. This is a more contentious exercise since it requires us to cardinalise health measures such that their effects on mortality can be added up and compared to each other. To this end, we follow the recently developed methodology outlined in Ferrer and Van Praag (2004). We take the line that mortality occurs when 'Latent Health' drops below a fixed threshold.[2] Taking this line, in Section 5 we define 'Latent Health' empirically as the factor that predicts death. Using this concept, we can add all the (short-term and long-term) effects of lifestyle choices, both directly and via the various other health measures, on mortality in a single cardinal measure. This allows us to talk about fractions of total effects of lifestyle choices on mortality that are captured (or not) by the other health measures and as such allows us to answer the question how well the effects of lifestyle on mortality are explained by other health measures. The answer we find can be summarised as 'not very well'.

Conclusions are drawn in Section 6.

2. A Short Review of the Literature

Over the last few decades, one of the most heavily researched topics in economics and the social sciences more generally is the determinants of health. Hundreds of studies have looked at the cross-sectional correlates of health, with only a few studies investigating the determinants of health in a dynamic setting using panel data. A general introduction to this vast field is in Jones (ed, 2007).

Attempts to understand different causal pathways through which socio-economic status and health affect each other have been numerous (see, for example, Smith, 1999 and Adler et al., 1994 for reviews). Pathways from health to wealth have been examined mainly based on the human capital theory by Grossman (1972), where health is assumed to be a stock that is

[2] This is partly motivated by the theoretical contributions of Grossman's health production model (2000, as adapted from 1972) and recent extensions to that model (Bolin et al. 2002; Eisenring 2000; Jacobson 2000). These papers have as their basic structure that individuals can invest time in order to defer the moment of death, thereby taking death to be the ultimate arbiter of whether an individuals underlying true health is good or bad. We adopt the same assumption, though the estimates in this paper cannot be seen as structural estimates of that model mainly because we include the possibility of there being orthogonal unanticipated factors in period-on-period health, whilst the Grossman model presumes individuals know exactly when they die and hence the absence of unanticipated orthogonal shocks.

built up through investment. Within this framework, health events can lead to considerable revisions of life-cycle decisions such as saving plans, labor supply, retirement expectations and bequest intentions (Smith, 2003). Pathways from wealth, or socio-economic status variables, to health have also been studied extensively in economics recently (Adams et al., 2003; Adda et. al., 2003; Hurd and Kapteyn, 2003; Meer et al., 2003; Michaud and Soest, 2004, Smith, 2003). Theories explaining a link between socio-economic variables and health have been put forward in these studies, but given the mixed nature of the results provided there does not appear to be a clear consensus emerging.

When examining the mechanisms by which health and socio-economic status are determined, it is important to realize that health and wealth are both dynamic processes that evolve over an individual's life-cycle. Health-relevant choices are made throughout the life-cycle and health outcomes affect the choice set at every moment in time. A comprehensive study of the co-evolution of these dynamic processes clearly requires a panel dataset with detailed information on lifestyle choices, major unforeseen events and relevant socio-economic variables.

From this perspective, the Health and Retirement Study (HRS) is a rich dataset which contains detailed health, wealth, income information together with lifestyle variables such as smoking, drinking, regular exercise, labour market dynamics, and a variety of marital and job outcomes. Given its richness, the HRS has recently been utilized extensively, not only to study health issues but also many other social and economic issues such as retirement and social security.

The HRS reports several different health indicators such as self-reported health, doctor diagnosed health problems, and mental health scores. In terms of measuring health, researchers have used these indicators in various forms. Adams et al. (2003) consider each of these health dimensions independently while they recognize that all indicators might be interrelated. Hurd and Kapteyn (2003) consider self-reported health status and Smith (2003) studies serious health conditions. Adda et. al. (2003) and Michaud and Soest (2004) build a health index combining some of those indicators into one dimension.

The HRS keeps track of mortality outcomes by recording deaths. Given its limited availability and inherent inability to allow for fixed-effects estimators (you die only once), however, mortality has not been studied as much as the other measures of health.

The general presumption across different disciplines, including economics, is that while smoking and heavy drinking are detrimental to health, regular exercise is beneficial. Having said this, there are studies that report opposite findings regarding drinking and exercise[3]. Using different datasets and approaches, the presumed consensus effects of smoking, drinking and other lifestyle behaviors and socio-economic characteristics on health and length of life have also been investigated empirically in numerous studies, but there still seems little agreement about the magnitudes and significance of each factor. (see, for example, Gardner and Oswald, 2004 for a detailed review of earlier literature on socio-economic correlates of mortality generally, and Sloan et. al., 2003 and 2004 for a detailed review of smoking and risk factor effects specifically).

3. Data and Definitions

3.1. Data

The data used in this paper are drawn from the first six waves of the Health and Retirement Study (HRS). The HRS is a nationally representative biannual panel for the US, which surveys approximately 7,600 households with a primary respondent between the ages of 51 and 61 during the first year of the study. The first wave of the panel was conducted in 1992, so the primary respondents represent cohorts born between 1931 and 1941 and our sample covers the period 1992-2002. If an age-eligible primary respondent had a spouse or partner co-residing then the spouse or partner was also given the same individual level interview separately even though he or she was not between the ages of 51 and 61. However, in collecting household level information, which would be the same for both spouses, only one interview is given generally to the financially responsible member of the household. In addition to a large number of usual demographic characteristics such as race, education and marital status, the survey collects detailed information on the nature of retirement decisions, expectations, housing, income and wealth holdings, work history, family composition, and the availability of insurance and pensions. Of particular interest for the present analysis is that the HRS provides detailed information on each respondent's health and cognitive status.

The HRS distinguishes death as a separate source of non-response. Our pooled sample has 58,422 year-person observations and an overall number of 1,248 deaths. The frequency of deaths in wave 2 through wave 6 is respectively: 216, 233, 239, 270, and 290.

[3] For drinking, we already mentioned the Sturm (2002) study that purports to find only negative effects from drinking. As to exercise, Quittan et. al. (1999) look at the impact of the impact of a three-month exercise program on the perception of quality of life in patients with chronic conditions and find only weak correlations between parameters of physical performance and quality of life domains. Since they evaluate an actual intervention rather than non-experimental data, their findings strongly suggest that the relation between exercise and health, especially in the short-run, is more complicated than previously thought.

3.2. Defining Health Status

We use five alternative but clearly related measures of health status or morbidity, and supplement this by also looking at mortality directly. These five measures are (i) self-assessed health, (ii) self-assessed difficulties with performing tasks, (iii) self-assessed mental health, (iv) doctor diagnosed long-term health conditions, and (v) Body Mass Index (BMI).

For the first measure, survey respondents are asked to rate their current health status on the familiar 1-5 scale, where 1 = Excellent Health (16.7% of observations), 2 = Very Good Health (29.8%), 3 = Good Health (29.5%), 4 = Fair Health (16.2%) and 5 = Poor Health (7.7%). Our second measure is a composite index relating to the level of difficulty the respondent has in performing a number of normal day-to-day activities or tasks. Specifically, we count the number of affirmative answers to the following 11 normal daily activities or tasks (percentage of observations in parentheses):

> *Whether the respondent has some difficulty with: Dressing (5.1%); Bathing (3.6%); Eating (1.4%); Getting in or out of bed (4.7%); Walking several blocks (19.2%); Sitting for two hours (18.6%); Getting up from a chair (30.0%); Climb several flights of stairs (33.4%); Lifting 10lbs (17.6%); Extending arms (12.1%); Pushing or pulling large objects (18.8%).*

where, for each question, the response takes a value of 1 if the answer is yes and 0 otherwise. Note that we have made no attempt here to place differential weights on these tasks. Similarly, our overall mental health variable aggregates answers to the questions:

> *Whether the respondent: Felt Depressed (13.5%); Everything was an effort (20.1%); Sleep was restless (26.4%); Was happy (86.1%); Felt Lonely (12.2%); Felt Sad (15.0%);, Could not get going (17.3%); Enjoyed life (91.5%).*

where, coding is 0 for No and 1 for Yes, and the responses to 4 and 8 are counted as positive and the rest as negative. Our fourth health measure uses information on doctor diagnosed health problems. In the HRS each individual is asked in each wave whether or not a doctor has ever told him or her that he or she had the following serious illnesses:

High blood pressure (41.1%); Diabetes (13.0%); Cancer (7.9%); Lung disease (7.8%); Heart problems (15.7%); Stroke (4.3%); Psychiatric problems (11.7%); Arthritis (46.8%).

Again, the coding takes a value of 0 for No and 1 for Yes, and our composite measure sums these variables. Finally, Body Mass Index (BMI) is the respondent's self-reported weight (in kilograms) divided by his or her square of height (in meters). Since the vast majority of the sample has a BMI of over 20, we will essentially take BMI as a linear variable (the issue that having a very low BMI is deemed unhealthy simply does not arise sufficiently often enough in the data). Interestingly, the average BMI in the HRS is approximately 27, which is clearly above what is regarded as a normal BMI measure by the US Department of Health and Human Services that ranges between 18.5 and 24.9 (25-29.9 is classified as overweight; 30+ is classified as obese).

3.3. Defining Health Investments

The clear advantage of the HRS for this study is that information is collected on a number of individual health investments, of which we focus on three. These are smoking, alcohol consumption and regular exercise.

Our smoking variable is simply a binary indicator of whether or not the respondent currently smokes. In the pooled panel, approximately 22 % of the sample fall into current smoker category. When we look at the data by each wave, it is notable that smoking rate is relatively higher in the first wave and there is a monotonically decreasing time trend over the sample period covered.

The questions relating to alcohol consumption asked in the HRS changed slightly between waves 1-2 and waves 3-6. To derive a consistent measure of drinking across all waves we create a drinking intensity variable which indicates the number of alcoholic drinks per week that the respondent consumes. In the pooled data, approximately 53 % of the individuals consume alcohol and the average number of standard drinks conditional drinking is 6.92.

Our regular exercise variable is also a binary indicator and shows whether the respondent participates in vigorous physical activity or exercise at least 3 times a week. However, the change in wording of the exercise questions between waves 1-2 and the rest of the data is slightly more problematic than for drinking. More explicitly, beginning from wave 3, the HRS asked each individual a single exercise question which is:

> *On average over the last 12 months have you participated in vigorous physical activity or exercise three times a week or more? By vigorous physical activity, we mean things like sports, heavy housework, or a job that involves physical labor.*

Thus, our dummy variable for waves 3 through 6 is set to 1 if the respondent's answered 1 to this question and 0 otherwise. For waves 1 and 2, on the other hand, we use answers to the following questions:

> Wave 1
>
> *Question 1: How often do you participate in vigorous physical exercise or sports -- such as aerobics, running, swimming, or bicycling? (Would you say 3 or more times a week, 1 or 2 times a week, 1 to 3 times a month, less than once a month, or never?)*
>
> *Question 2: How often do you do heavy housework like scrubbing floors or washing windows?*
>
> Wave 2
>
> *How often do you participate in vigorous physical activity or sports - -such as heavy housework, aerobics, running, swimming, or bicycling?*

For wave 1, the derivation of our exercise dummy combines those two questions. The physical exercise indictor is set to 1 if either is three times or more a week. For wave 2, the derivation is based on the number of times and frequencies to arrive at a yes or no conclusion. The dummy is set to 1 if the period is day, or if the respondent answers 3+ times a week, 12+ times a month, or 156+ times a year. Given our definition, we observe that in the pooled panel about 37 % of the population exercise regularly at least three times or more per week.

4. What Do Ordinal Models Tell Us?

In this Section we attempt to be as close to the current health economics literature as we can by looking at the robustness of very standard models. As such, we recognize that self-assessed health, self-assessed difficulties with performing tasks, self-assessed mental health, and doctor diagnosed long-term health conditions are ordinal in nature.

As our first model we will use the recently developed conditional fixed effects logit model (Ferrer and Frijters 2004), which has recently been used to look at the effect of income on health satisfaction in the German Socio Economic Panel (Frijters et al. 2005). The model is of the form:

$$\begin{aligned} H_{it}^{*} &= x_{it}\beta + \delta_t + f_i + \varepsilon_{it} \\ H_{it} &= k \Leftrightarrow H_{it}^{*} \in [\lambda_k, \lambda_{k+1}] \end{aligned} \tag{1}$$

where H_{it}^{*} is latent health corresponding to one of the self-assessed measures in our data; H_{it} is an observed ordinal indicator of health; λ_k is the cut-off point (increasing in k) for the attitudinal answers; x_{it} is observable time-varying characteristics; δ_t denotes time-varying general circumstances; f_i is an individual fixed characteristic; and ε_{it} is a time-varying logit-distributed error-term that is orthogonal to all characteristics.

Our conditional estimator for δ_t and β maximizes the following conditional likelihood:

$$\begin{aligned} &L[I(H_{i1} > k_i), .., I(H_{iT} > k_i) \mid \sum_t I(\mathrm{H}_{it} > k_i) = c] \\ &= \frac{e^{\sum_{t=1}^{T} I(H_{it} > k_i) x_{it}\beta}}{\sum_{H \in S(k_i, c)} e^{\sum_{t=1}^{T} I(\mathrm{H}_{it} > k_i) x_{it}\beta}} \end{aligned} \tag{2}$$

which is the likelihood of observing which of the *T* stated health outcomes of the same individual are above k_i, given that there are *c* out of the *T* outcomes that are above k_i. Here, $S(k_i, c)$ denotes the set of all possible combinations of $\{H_{i1}, .., H_{iT}\}$ such that $\sum_t I(\mathrm{H}_{it} > k_i) = c$. Also, H_{it} is used to denote the random variable and H_{it} the realization.

As we see, the fixed-effects have dropped out of this likelihood. It therefore yields estimates only for δ_t and β. This model is an extension of the fixed-effect logit model by Chamberlain (1980). Unlike the Chamberlain methodology that recodes the data such that only a crossing over a barrier that is the same for everyone (say, *k*) can be used, our model uses crossings over person specific barriers (say, k_i). When some individuals for instance only report values between 'bad' and 'very bad', and others only between 'good' and 'very good', then using the same barrier for everyone cannot record changes for both groups of individuals. Those individuals then have to be dropped from the estimation procedure. With individual specific barriers all individuals whose health outcomes differ over time, can be included. The most important advantage is therefore that it allows us to use the vast proportion of the observations. The model is estimated by Maximum Likelihood in GAUSS.

For mortality we run a simple Probit on whether an individual dies in the next period, using the same set of explanatory variables as used for the ordinal measures. Because BMI is not an ordinal variable but a continuous one, we estimate its determinants by Generalized Least Squares.

Empirical Results

In all specifications, it was the case that we started with initial specifications that included many more marital, wealth, and income variables. We decided in final specifications to only use total income and total wealth though because of the fear of high measurement error in specific items of wealth and income (such as government income and housing wealth). The exclusion of more detailed income and wealth results did not significantly affect any of the other coefficients.

The full specifications are in the Appendix. Here we focus only on the effects of age, lifestyle choices, income, and wealth on the 6 different health measures:

Table 1: The Effect of Choice Variables, Age and Income on Ordinal Health Outcomes in the HRS Data 1992-2002

	Stated Health	**Difficulties with Tasks**	**Mental Health**	**BMI**	**Serious Illnesses**	**Death**
Smoking at t	0.221**	-0.400**	0.155	-0.565	0.343	0.031
No. of drinks at t	0.000	0.000	-0.004	0.008**	-0.015	-0.006
Regular exercise at t	0.344**	-0.332**	0.114**	-0.152**	-0.193	-0.405**
Log income at t	0.004	0.040**	0.014	0.028**	0.136	-0.032**
Smoking at t-3	-0.028	-0.184*	0.052	0.024	-0.841	0.299**
No. of drinks at t-3	-0.006	0.000	-0.008	-0.005	-0.031	-0.004
Regular exercise at t-3	-0.051	0.193**	0.053	0.022	0.131	-0.024
Log income at t-3	-0.011	-0.003	0.020	0.001	-0.107	-0.024*
Age	0.018	0.055**	0.028**	0.007	7.022**	0.035**

* indicates significant at 90%; ** indicates significant at 95%.

From Table 1 we see that the effects of various life styles, age, and income differ greatly for the various health measures and mortality. Take exercise. It improves all health indicators in the short-run, i.e. reduces the risk of mortality, improves self-stated and doctor-assessed health, decreases BMI, reduces the number of difficulties with tasks and improves mental health. The long-run effects are contradictory though: the effect of exercise in 3 period past (which is 6 years ago) is insignificant for all health measures except for self-stated health. Furthermore, the point-estimate of the long-term effect is to increase the difficulties with tasks and to increase BMI, as well to have no lasting effect on mortality (i.e. in order to have a positive effect on mortality, one would have to keep up an exercise regime). Next, look at income. Whilst it reduces mortality both in the long-run and in the short run (by a relatively small amount though: an increase in income by 10% would reduce the probability of death by about 0.00015), a higher income is related to more difficulties with tasks and has no significant effect on self-stated health. The long-run effect of income is insignificant for all non-mortality health measures. Perhaps most striking is the effect of smoking. Smoking has a strong mortality increasing effect in the long-run. Yet, the long-run effect of smoking on the number of serious illnesses is decreasing, and the effects on self-stated health are found to be positive in the short run and insignificant in the long-run. Whilst one could argue selection effects cause the unanticipated short-run positive effect on self-stated health, it is nevertheless indicative of a lack of robustness in the effect of smoking on the various health measures. One could argue that a health measure that fails to pick up the negative effects of smoking, even whilst controlling for fixed effects, must be viewed with suspicion.

The general unimportance of drinks for all of these health measures except BMI is noteworthy. We re-ran all equations adding the square of the number of drinks to see whether drinks would have an effect at very high or low levels. We however found no significant non-linearity in any of the specifications (neither in the current number of drinks or the lags thereof). This is probably due to the low number of individuals who drink a lot in this sample: only 1.5% of the sample drink more than 3 alcoholic consumptions per day, only 0.7% drink more than 4 consumptions per day, and only 0.01% drink more than 6 consumptions per day. The age range in our sample probably does not cover the main binge drinkers of the population.

5. What a Cardinalisation of the Health Measures Tells Us

Here, we depart from the assumption of ordinality because we cannot compare the effects of lifestyles on different ordinal outcomes with each other, nor add them up in terms of their effects on mortality. We will thus treat each non-mortality health measure as cardinal. Our cardinal model extends the basic descriptive methodology of Frijters (1999, Chapter 5), Van Praag et al. (2003), and Ferrer and Van Praag (2004) to allow for fixed-effects and to allow for an outcome variable that is observed at most once:

$$LH_{it}^{*} = \sum_{j} \alpha_j H_{it}^{j} + \sum_{s} \alpha_{js} H_{it-s}^{j} + X_{it}'\beta + \sum_{s} X_{it-s}'\beta_s + \varepsilon_{it}$$

$$\text{Death at } t+1 \rightarrow LH_{it}^{*} < 0, \text{otherwise } LH_{it}^{*} \geq 0 \qquad (3)$$

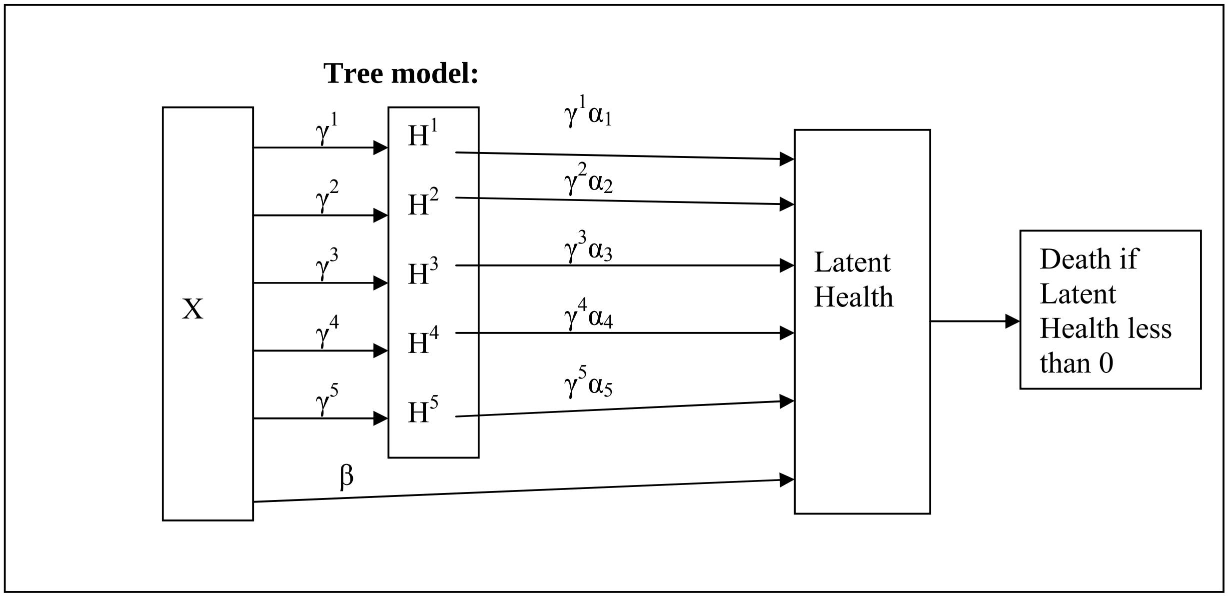
$$H_{it}^{j} = X_{it}'\gamma_j + \sum_{s} X_{it-s}'\gamma_j^{s} + v_i^{j} + \mu_{it}^{j}$$

where X_{it} is a set of individual and household characteristics that includes the health investments and life circumstances; LH_{it} denotes 'Latent Health', which is in the short run affected by X_{it} directly (via $X_{it}'\beta$), and indirectly (the terms $X_{it}'\gamma_j$ that appear in Latent Health via $\sum_{j} \alpha_j H_{it}^{j}$); We cannot actually observe latent health in any period but can only infer something about its level from whether a person lives or dies in a period; H_{it}^{j} (for j=1,..,5) denotes a particular observed health variable; v_i^{j} are fixed-effects that denote initial characteristics that can correlate with X_{it} in an arbitrary manner. ε_{it} are i.i.d. normally distributed shocks; μ_{it}^{j} are independently distributed shocks that are unrelated to X_{it}; the long-run effects of X_{it} are reflected in the terms $\sum_{s} X_{it-s}'\beta_s$ and $\sum_{s} X_{it-s}'\gamma_j^{s}$. The long-run effects of the health variables are reflected in $\sum_{s} \alpha_{js} H_{it\text{-}s}^{j}$.

Our innovation with respect to the previous work with this descriptive model is twofold. Firstly, our model specifies the paths via which individual characteristics X_{it} affect Latent Health. As such, it gives detail to how and why a particular variable has the effect on mortality that it has. This means our model is able to ascertain how much the observed health variables H_{it}^{j} actually capture the effects of X_{it} and thus how good a proxy those variables are for Latent Health. By conceptualizing Latent Health as the factor that predicts death, we also tie our analysis into the large theoretical literature that has taken health investments to mean investments in deferring the moment of death. Secondly, the fixed-effect terms in the health equations for H_{it}^{j} improve upon most previous work that has assumed the absence of individual fixed effects that correlate with X.

The short-run effects of X in our modeling framework are summarized in Figure 1.

Figure 1: The Short-run Direct Effects and the Short-run Indirect Effects of X



Based on this model, we can now define the following statistics of interest:

Short-run, direct effect of ΔX: β

Long-run, direct effect of ΔX: $\beta + \sum_s \beta_s$

Short-run, indirect effect of ΔX: $\sum_j \alpha_j \gamma_j$

Long-run, indirect effect of ΔX: $\sum_j [(\alpha_j + \sum_s \alpha_j^s)(\gamma_j + \sum_s \gamma_j^s)]$

Here, a short-run effect refers to the effect that a change in *X* has in the year of the change. A long-run effect refers to the effect that a permanent change in *X* has, where permanent is empirically defined as at least as long as 3 waves, which is equivalent to 6 years. One can interpret these long-run effects as the cumulative effects of a variable. The ratio of total indirect to total effects (direct + indirect) is then a measure of the extent to which the health variables H_{it}^j actually capture the Latent Health effects of X_{it}.

The simple steps we follow to estimate the model are:

Step 1: Run Generalised Least Squares (GLS) fixed-effects to obtain estimates of each H_{it}^j separately; we therefore obtain the γ parameters.

Step 2: Run a Probit analysis on mortality; we therefore obtain the α and β parameters.

The intuition for the use of our linear model is that it allows the simultaneous inclusion of a mass of information, whilst also affording fairly simple interpretations of parameters and yet retains the ability to use fixed effects. We could not achieve this without the cardinalisation imposed above. When one instead would, for instance, include all the intermediate health outcomes as a whole vector of binary outcomes (eg. whether health satisfaction is 7 or not; whether health satisfaction is 8 or not; whether each of the mental health components is 1 or not, etc.) then one looses the ability to include fixed effects and also looses the ability to give simple interpretations to the findings. We may here shortly expand on the disadvantages of the main alternatives. One alternative is to replace Step 2 with a survival analysis. The issue with that is that our data is not well suited for this because only the interval of time of death is known. The other main alternative is to replace the use of H_{it}^{j} as continuous variables in both Step 1 and 2 by an ordinal estimator. The 'loss' of such changes to the model used is that those options do not allow us to take account of fixed effects in the health indicators: the main ordinal models with fixed effects (probits and logits) are either biased for small numbers of observations per individual (probit), or only allow for conditional estimators from which one cannot back out the actual levels of the latent variables (logit). If one is not even prepared to make the parametric assumptions underlying the probit and logit models and instead wants to have semi-parametric estimators with unspecified distributions of the error terms of ordinal intermediate health outcomes, then the ability to back out the level of the latent variable in the second step disappears completely. Therefore, our linear model should be seen as a workable first-order approximation which can be readily used by a wide range of researchers.

Empirical Results

We will begin by discussing the estimated determinants of H_{it}^{j}, followed by the determinants of Latent Health. For all 5 measures of health, it held that the effects of demographic and income variables were very small, with only now and then significant results, such as a negative effect of the number of years worked and longest tenure on the number of serious illnesses. Interestingly, income's only significant effect is to reduce BMI. The only significant effect of the number of divorces is to strongly increase BMI. There are strong differentials to be found amongst various occupations (these occupations are the ones the respondent worked in 'most of the time'), with the default (no job) being the least healthy occupation for nearly all health measures. In the Appendix we have reported the results on each of the 5 health measures.

We focus now on the short-run and long-run coefficients of 5 life choices and life circumstances of interest, where an asteric denotes significant at the 95% confidence level:

Table 2: The Effect of Choice Variables and Age on Intermediate Health Outcomes in the HRS Data 1992-2002

	Short-Run				
	Self-Assessed Health	**Difficulties with Tasks**	**Mental Health**	**BMI**	**Serious Illnesses**
Smoking	0.093762*	-0.318620*	0.113084	-0.564680*	-0.178980*
No. of drinks	0.000260	-0.002760	-0.004290	0.007995*	-0.003930*
Regular exercise	0.120015	-0.271270*	0.116243*	-0.152480*	-0.031570*
Log income	0.000575	0.016280	0.018361	0.028172*	0.000751
Age	0.001649	0.049520*	0.017345*	0.007255	0.119687*
	Long-Run				
Smoking	0.075560*	-0.435190*	0.153794*	-0.540810*	-0.197170*
No. of drinks	-0.002090	0.000857	-0.009040	0.003240	-0.002020
Regular exercise	0.104637	-0.196120*	0.122661*	-0.130210*	-0.029810*
Log income	-0.001460	0.007976	0.024988*	0.0292340*	-0.001390
Age	0.075560*	-0.435190*	0.153794*	-0.540810*	-0.197170*

We can see several surprising results here. The main one is that the short-run and the long-run coefficients of smoking are all positive. A person who smokes has, in the year after she takes up smoking (i.e. the short-run effect), a self-assessed health that is 0.094 higher than an otherwise identical person who doesn't smoke. That person has a BMI that is 0.56 lower in the year after taking up smoking; and a mental health that is 0.11 higher. In the short run hence, smoking increases stated health, reduces the problems with chores, improves mental health, reduces BMI very strongly, and reduces the number of serious illnesses measured. One objection one could have is that there may be simultaneity in these measures: perhaps some unobserved improvement in each of these measures correlates with deciding to become a smoker again. In such a case, however, we should have found the long-run effects to be very different from the short-run effects. This is not the case. Indeed, the long-run 'benefits' of smoking are stronger than the short-run benefits for problems with chores, mental health, and the number of serious illnesses. As we will see later, the known negative effect of smoking in terms of a much higher chance of premature death is found in our data also, but appears not to be picked up by the health measures above, which means they fail to pick up the deleterious effects of smoking.

Another surprise is the unimportance of drinking: the short-run effect of drinking is not only small in all cases, but only significant for serious illnesses (decreasing) and BMI (increasing).

There is no long-run significant effect of the number of drinks on any of these 5 health outcomes. Again, this is probably due to a lack of variation at the high-number of drink end in this data.

Frequent exercise is on the other hand always associated with better health and significantly so except for stated health. Moreover, the long-run effect and the short-run effect are basically the same for all health outcomes. This means that only a sustained regime of frequent exercise brings a sustained health improvement: any exercise done more than two years ago has no significant effect anymore on any of these health outcomes.

With income, we find short-run effects that are very different from long-run effects. In the short-run, income increases relate to more problems with chores, improved mental health, and more serious illnesses. In the long-run, income increases have an insignificant negative effect on the number of serious illnesses, whilst there is a long-run positive effect on mental health and BMI. There is no significant effect on stated health, in line with other studies that find very little income effects (see Adams et al., 2003, and Meer et al., 2003, for US evidence).

For the 5 health variables we have the following short-run and long-run relations with Latent Health, where a higher Latent Health means a lower probability of dying next period:

Table 3: The Short-run and Long-run Effect of Intermediate Health Variables on Latent Health in the HRS data 1992-2002

	Effects on Latent Health	
	Total SR	**Total LR**
Self-assessed health status	0.216*	0.296*
Difficulties with tasks	-0.039*	-0.013
Mental health	0.011	-0.067*
BMI	0.078*	0.022
Serious illness	-0.247*	-0.145*

Table 3 confirms that stated health is a strong predictor of mortality, and that the majority of the total information in stated health (as reflected by the long-run effect) is already in the last observed stated health (the short-run effect). Difficulties with tasks have a slight negative effect on Latent Health, but not a significant long-run effect implying that all the information in task difficulties is subsumed by the other health measures. In the short-run, increases in BMI are actually health improving, though long-run effects are again insignificant (which will be mainly due to the fact that the other health measures already capture the negative effects of BMI). Serious illnesses have a strong negative effect on Latent Health, although the long-run effects are less bad than the short-run effects, rather like the flu is worse in the short-run than in the long-run. Mental health interestingly enough has a negative long-run effect on Latent Health. In summary, stated health and serious illnesses are the main health variables that have plausible and strong effects on Latent Health.

Next, we turn to the effects of the choice variables on Latent Health. The statistics of interest on these variables read (where ME means marginal effect on probability of death in a 2-year period, evaluated at the mean of the sample):

Table 4: The Short-run, Long-run, Direct, and Indirect Effect of Choice Variables on Latent Health in the HRS Data 1992-2002

	Total SR	**ME**	**Total LR**	**ME**	**Direct SR**	**Indirect SR**	**Direct LR**	**Indirect LR**
Smoking	-0.018*	-0.0006	-0.222*	-0.0076	-0.052	0.034*	-0.257*	0.035*
No. of drinks	0.007*	0.0002	0.003	0.0001	0.005	0.002	0.003	0.000
Regular exercise	0.255*	0.0084	0.185*	0.0061	0.222*	0.034*	0.158*	0.027*
Log income	-0.005	-0.0001	0.022	0.0007	-0.006	0.002	0.023	-0.001
Total wealth $m	0.063	0.0021	0.089	0.0029	0.063	0.000	0.089	0.000
Age	-0.067*	-0.0022	-0.055*	-0.0018	-0.036*	-0.030*	-0.036*	-0.019*

Table 4 forms the bottom line of our inquiry into cardinality. We see strong negative long-run effects of smoking on Latent Health, where the short-run effect is much smaller than the long-run effect. We find that the indirect effects of smoking are opposite to the direct effects, implying that the observed health variables fail to pick up the actual Latent Health effects of smoking. We may even in this context remark on the health/death trade-off that smoking induces: smoking improves stated health by about 0.08 each year, whilst it increases the probability of dying the next 2 years by 0.8% (seen in the Table by the marginal long-run effects). Translated into our 10 year period, this means about an expected 3 months less life (=0.008*(2+4+6+8+10) years) is traded in for 0.8 higher stated health in one year, or in other words having an improvement from fair to good health for one year is worth about 4 months of life to a person.

Drinking is surprisingly unimportant for Latent Health, both in the short-run and in the long-run. Although one may rightfully wonder whether very heavy drinking does have negative effects (our data is not really suited to answer that question because the vast majority of the time-variation in our sample is between grades of moderate drinking), we would venture that the complete absence of significant effects on this linear model should make one sceptical of finding much with other specifications.

Frequent exercise is found to have strong positive effects, both in the short-run, the long-run, directly, and indirectly. About 20% (=0.027/(0.027+0.15*)) of the long-run positive effect of frequent exercise runs via the 5 health measures. The effect of frequent exercise is almost of the same magnitude as smoking, implying that the combined effect of starting to smoke and doing frequent exercise will in the long run have no Latent Health effects, though such a combination would have short-run Latent Health improvements because the negative effects of smoking are long-run whereas exercise has strong short-run benefits.

Income and wealth have surprisingly little effect of their own (taking the other choices as given): one would need about 2.5 million dollars to offset the long-run effect of smoking.

The effect of age is, not surprisingly, strongly negative on Latent Health, where we may see that about 1/3 of the negative long-run effect of age is picked up via the 5 health measures.

6. Conclusion

In this paper we investigated the robustness of the determinants of various health measures using the HRS that has followed individuals biannually from 1992 to 2002. Robustness here refers to whether the effects of a particular variable on mortality is also picked up by 5 other health measures (self-stated health, doctor assessed health, mental health, BMI, and self-stated difficulty with chores) in a consistent manner, and robustness also refers to whether the short-run effects of variables are the same as long-run effects.

The lifestyle choice for which the results are least robust across health measures and hence for which one cannot generalise across health measures is smoking. We found the expected result that smoking has a strong increasing long-run effect on mortality, though we find no short-run effect on mortality (which may be due to the fact that people about to die don't take up smoking again). Surprisingly though, smoking improves self-stated health (both in the short and in the long-run which means it is not due to a simultaneity problem), reduces the problems individuals have with chores, improves mental health, reduces the number of measured serious illnesses, and strongly reduces Body Mass Index (BMI). The indirect effects of smoking on mortality via these other health measures is thus opposite to the direct effect. Nevertheless, the direct long-run effect on mortality strongly outweighs all these indirect effects. This means that the stated and measured health problems insufficiently capture 'Latent Health', and that there can be large differences in short-run, long-run, direct, and indirect effects across health measures.

The discrepancies are less pronounced for other life style choices, but there are still some strong discrepancies to be found. We find that physical exercise significantly reduces mortality both in the short-run and in the long-run, whereby only 20% of that effect is picked up via the health measures we used. The effect of regular exercise is not significant though for self-assessed health whereas it is significant for all other measures of health, including mortality. We find no evidence for any long-run effect of drinking on any of the health measures. We do find a significant short-run reduction in the number of serious illnesses when drinking more, which is not found for any of the other health measures implying that for drinking too, the effects across measures do not necessarily carry over. We find no short-run or long-run benefits of income or wealth on mortality. This does not necessarily preclude the possibility that income and wealth has an effect though, but simply means that the effect of wealth and income would have to work via the increased levels of exercise and reduced levels of smoking associated with higher income/wealth levels rather than via channels not identifiable in this data (which is what the direct coefficients pick up). Importantly for the issue of comparability though is that the effects of income and wealth are once again not uniform across health measures. Income for instance significantly increases BMI and improves mental health in the long-run, yet has no significant effect on any of the other measures. Indeed, income increases the number of difficulties with tasks though not significantly so.

We thus do not find evidence for robustness of effects of life style variables across different health measures and instead find significant discrepancies across measures, especially for smoking. We furthermore find that even when health measures are in agreement about the sign, such as for exercise, the cumulative effect of exercise on mortality via the other health measures is only a small proportion of the entire effect of exercise on mortality, implying that health measures used apparently pick up only a small fraction of the factors important for mortality. All this makes it important for future analyses to reduce the value put on stated health, BMI, mental health, difficulties with chores, and even on doctor-assessed major illnesses: they do not fully capture the factors responsible for mortality, even though they are themselves strongly and plausibly related to mortality (as also found in other studies, like Idler and Kasl (1995) and Idler and Benyismini (1997)). The different health measures furthermore are affected in quite different ways by different lifestyle variables, implying that results on any one of them do not carry over to any of the other ones.

References

Adams, P., Hurd, M., McFadden, D., Merrill, A. and Ribeiro, T. (2003), 'Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status', *Journal of Econometrics*, 112, pp. 3-56.

Adda, J., Tarani, C., and Marmot, M. (2003), 'Socioeconomic status and health: causality and pathways', *Journal of Econometrics*, vol. 112, no 1, pp. 57-63.

Adler, N. *et al*. (1994), 'Socioeconomic status and health, the challenge of the gradient', *American Psychologist*, 49, pp. 15-24.

Attanasio, O. and Hoynes, H. (2000), 'Differential mortality and wealth accumulation', *Journal of Human Resources*, 35, pp. 1-29.

Benzeval, M., Taylor, J. and Judge, K. (2000), 'Evidence on the relationship between low income and poor health: Is the government doing enough?', *Fiscal Studies*, 21, pp. 375-399.

Benzeval, M. and Judge, K. (2002), 'Income and health: The time dimension', *Social Science and Medicine*, forthcoming.

Bolin, K. Jacobson, L., Lindgren, B. (2002) 'The Family as the Health Producer-When Spouses Act Strategically', *Journal of Health Economics*, vol. 21, no. 3, pp. 475-95

Burstrom, B. and Fredlund, P. (2001), 'Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes?', *Journal of Epidemiology and Community Health*, 55, pp. 836-840.

Caruso, L., Silliman, R., Demissie, S., Greenfield, S., and Wagner, E. (2000), 'What Can We Do to Improve Physical Function in Older Persons With Type 2 Diabetes?', *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 55:M372-M377.

Case, A., Fertig, A. and Paxson, C. (2004), 'The lasting impact of childhood health and circumstances', Center for Health and Wellbeing Discussion Paper, Princeton.

Chamberlain, G. (1980), 'Analysis of covariance with qualitative data', *Review of Economic Studies*, vol. 47, pp. 225-238.

Cutler, D. and Glaeser, E. (2005), 'What explains differences in smoking, drinking, and other health-related behaviours?' Working Paper, Harvard University.

Di Castelnuovo, A., Costanzo, S., Bagnardi, V., Donati, MB., Iacoviello, L., and De Gaetano, G. (2006), 'Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies', *Archives of Internal Medicine*, December 11-25:166 (22): 2437-2445.

Eisenring, C. (2000), 'Is There a Trade-Off between Longevity and Quality of Life in Grossman's Pure Investment Model?' *Health Economics*, vol. 9, no. 8, pp. 669-80.

Ezzati, M. and A.D. Lopez (2003), 'Estimates of Global Mortality Attributable to Smoking in 2000' *Lancet*, vol. 362, pp. 847-852.

Ferrer-i-Carbonel, A., Frijters, P. (2004), 'The effect of methodology on the determinants of happiness', *Economic Journal* 114, 641-659

Ferrer-i-Carbonel, A., Van Praag, B.M.S. (2004), *Happiness Quantified: a satisfaction calculus approach*, Oxford University Press.

Frijters, P. (1999), *Explorations of welfare and well-being*, Tinbergen Insitute Thesis, Thela Thesis: Amsterdam.

Frijters, P., Shields, M.A., and J.P. Haisken-DeNew (2005), 'The effect of income on health: evidence from a large scale natural experiment', *Journal of Health Economics* (24), pp. 997-1017. IZA working paper.

Gardner, J. and Oswald, A. (2004), 'How is mortality affected by money, marriage, and stress?', *Journal of Health Economics*.

Gerdtham, U. and Johannesson, M. (2004), 'Absolute income, relative income, income inequality, and mortality', *Journal of Human Resources*, 39, pp. 228-248.

Grossman, M. (1972), 'On the Concept of Health Capital and the Demand for Health', *Journal of Political Economy,* 82, 223-255.

Grossman, Michael (2000), 'The Human Capital Model', *Handbook of health economics*. Volume 1A. 2000, pp. 347-408

Hurd, M. and Kapteyn, A. (2003), 'Health, wealth, and the role of institutions', *Journal of Human Resources*, vol. 83, no. 2, pp. 386-415.

Idler, E. and Benysmini, Y. (1997), 'Self-reported health and mortality: A review of twenty-seven community studies', *Journal of Health and Social Behavior*, 38, pp. 21-37.

Idler, E. and Kasl, S. (1995), 'Self-rating of health: Do they also predict change in functional ability?', *Journal of Gerontology: Social Sciences*, 508, pp. S344-S353.

Jacobson, L. (2000), 'The Family as Producer of Health-An Extended Grossman Model', *Journal of Health Economics*, vol. 19, no. 5, pp. 611-37.

Jones AM (2007, ed). Elgar *Companion to Health Economics*. Edward Elgar. Forthcoming

Jones AM, X Koolman and N Rice, (2004) 'Health-related attrition in the BHPS and ECHP: using inverse probability weighted estimators in nonlinear models', ECuity III Working Paper 18

Knupfer, G. (1987), 'Drinking for Health: the daily light drinker fiction', *Addiction*, 82 (5).

Liao, Y., McGee, D.L., Cao, G., and Cooper, R.S. (2000), 'Alcohol intake and mortality: findings from the National Health Interview Surveys (1988 and 1990)', *American Journal of Epidemiology*, April 1:151 (7), 651-659.

Meer, J., Miller, D. and Rosen, H. (2003), 'Exploring the health-wealth nexus', *Journal of Health Economics*, 22, pp. 713-730.

Michaud, P. and Soest, A. V. (2004), ‘Health and wealth of elderly couples: causality tests using dynamic panel data models’, IZA Discussion Paper No. 1312.

Quittan M., Sturm B., Wiesinger G. F., Pacher R., and Fialka-Moser V. (1999), ‘ Quality of life in patients with chronic heart failure: A randomized controlled trial of changes induced by a regular exercise program’, *Scandinavian Journal of Rehabilitation Medicine*, 31 (4), 223-228.

Ruhm, C. (2000), ‘Are recessions good for your health?’ *Quarterly Journal of Economics*, 115, pp. 617-650.

San José, B., Van Oers, J.A.M., Van De Mheen, H., Garretsen, H.F.L., and MacKenbach, J.P. (2000), ‘Drinking patterns and health outcomes: occasional versus regular drinking’, *Addiction* 95 (6), 865–872.

Sloan, F., Smith, K., and Taylor, D. (2003), ‘The smoking puzzle: Information, risk perception, and choice’, Harvard University Press.

Sloan, F., Osterman, J., Conover, C., Taylor, D., and Picone, G. (2004), ‘The price of smoking’, MIT press.

Smith, J. (1999), ‘Healthy bodies and thick wallets: The dual relation between health and economic status’, *Journal of Economic Perspectives*, 13, pp. 145-166.

Smith, J. (2003), ‘Consequences and predictors of new health events’, NBER Working Papers 10063.

Sturm, R. (2002), ‘The effects of obesity, smoking, and drinking on medical problems and costs’, *Health Affairs*, March/April, pp. 245-253.

US Department of Health and Human Services (1989), ‘Reducing health consequences of smoking: 25 years of progress’, A Report of Surgeon General, DHHS Publication No. (CDC) 89-8411

Van Praag, B.M.S., Frijters, P. , Ferrer-i-Carbonel, A., (2003), ‘The anatomy of subjective well-being’, *Journal of Economic Behavior and Organisation*, 51, pp. 29-49.

Appendix

Table 5: Conditional Logit Fixed-Effect Estimates of the Determinants of Various Health Status Measures

	Self-Assessed Health		**Difficulties with Chores**		**Mental Health**		**Serious illness**	
	Coeff.	**|*t*|**	**Coeff.**	**|*t*|**	**Coeff.**	**|*t*|**	**Coeff.**	**|*t*|**
Age	0.018	1.47	0.055	5.00	0.028	2.47	7.022	3.06
No. of times married	-0.204	0.93	-0.038	0.16	0.040	0.19	2.672	0.01
No. of times divorced	-0.082	0.32	-0.574	2.22	-0.280	1.22	4.041	0.01
Regular exercise at *t*	0.344	7.85	-0.332	7.81	0.114	2.73	-0.193	0.04
Regular exercise at *t*-3	-0.051	1.19	0.193	4.57	0.053	1.25	0.131	0.03
No. of drinks at *t*	0.000	0.03	0.003	0.64	-0.004	0.96	-0.015	0.04
No. of drinks at *t*-3	-0.006	0.79	0.000	0.02	-0.008	1.25	-0.031	0.04
Smoker at *t*	0.221	2.25	-0.400	4.11	0.155	1.59	0.343	0.03
Smoker at *t*-3	-0.028	0.28	-0.184	1.95	0.052	0.56	-0.841	0.07
Spouse's health	0.122	4.60	-0.042	1.62	0.158	6.22	0.192	0.07
No. of children	0.004	0.22	0.040	2.27	0.014	0.85	0.136	0.06
No. of siblings	-0.011	0.73	-0.003	0.25	0.020	1.44	-0.107	0.10
Log of household income at *t*	-0.074	1.67	0.036	0.83	0.050	1.18	-0.036	0.01
Log of household income at *t*-3	-0.007	0.46	0.026	1.74	-0.013	0.87	-0.067	0.05
Years of tenure in longest reported job	0.011	0.89	0.005	0.42	-0.006	0.49	0.214	0.02
Total years of work	0.004	0.21	0.034	1.78	0.047	2.49	0.899	0.11
Managerial Oper.	0.138	1.06	-0.329	2.54	-0.066	0.53	-0.459	0.03
Profess. Operator	0.480	3.54	-0.304	2.22	0.236	1.79	0.789	0.03

Sales	0.494	3.74	-0.479	3.76	0.197	1.57	0.702	0.04
Clerical/admin	0.299	2.68	-0.224	1.99	0.098	0.89	-0.846	0.02
Services: cleaning	0.130	0.48	-0.138	0.57	0.238	0.94	1.809	0.05
Services: security	1.095	3.32	-1.099	3.49	-0.043	0.13	-2.708	0.08
Services: food prep.	0.384	1.49	-0.137	0.61	0.072	0.32	-2.646	0.11
Health services	0.247	1.01	-0.533	2.20	-0.060	0.25	1.089	0.03
Personal services	0.299	2.01	-0.011	0.07	0.089	0.61	-0.875	0.05
Farming/fishing/forest	0.452	1.97	-0.273	1.15	0.102	0.42	1.501	0.04
Mechanics/repair	0.430	1.70	-0.206	0.80	-0.224	0.86	1.970	0.09
Construction	-0.011	0.05	-0.325	1.31	0.188	0.74	0.045	0.00
Precision production	0.600	2.57	-0.426	1.81	0.217	0.88	-2.287	0.04
Machine operator	0.120	0.67	-0.313	1.71	0.240	1.43	-1.434	0.03
Transport operator	0.616	3.22	-0.688	3.65	0.212	1.12	-2.889	0.12
Other operators	0.446	1.85	-0.419	1.72	0.267	1.09	2.469	0.16
Mean Log L	-1.06072		-1.04349		-1.06419		-0.003	
Number of observations	15962		17487		17230		9084	
Number of individuals	5425		5963		5864		3092	

Table 6: Probit Model of Mortality without Health Measures

	Coeff.	$\|t\|$	**ME**
Age	0.035	8.39	0.0016
No. of times married	0.056	1.29	0.0020
No. of times divorced	0.059	1.31	0.0021
Regular exercise at *t*	-0.405	8.86	-8.8600
Regular exercise at *t*-3	-0.024	0.43	0.0025
No. of drinks at *t*	-0.006	1.39	0.0002
No. of drinks at *t*-3	-0.004	0.82	0.0002
Smoker at *t*	0.031	0.48	0.4800
Smoker at *t*-3	0.299	4.95	4.9500
Spouse's health	-0.035	2.07	0.0008
No. of children	-0.032	2.13	-0.0015
No. of siblings	-0.024	1.94	0.0006
Log of household income at *t*	-0.017	1.83	-0.0008
Log of household income at *t*-3	-0.019	2.21	-0.0009
Years of tenure in longest reported job	-0.002	1.19	-0.0001
Total years of work	0.003	1.76	0.0001
Managerial Oper.	-0.443	3.76	0.0024
Profess. Operator	-0.534	4.58	0.0021
Sales	-0.345	3.06	0.0027
Clerical/admin	-0.691	5.56	0.0018
Services: cleaning	-0.210	0.84	0.0074
Services: security	-0.502	1.77	0.0043

Services: food prep.	-0.772	2.81	0.0023
Health services	-0.314	1.53	0.0049
Personal services	-0.570	3.50	0.0024
Farming/fishing/forest	-0.252	1.22	0.0056
Mechanics/repair	-0.030	0.17	0.0075
Construction	-0.360	1.56	0.0049
Precision production	-0.204	1.05	0.0059
Machine operator	-0.194	1.38	0.0044
Transport operator	-0.528	2.94	0.0027
Other operators	-0.456	1.68	0.0046
Constant	-3.261	10.48	0.0000
Log L	-2158.17		
Pseudo R2	0.1842		
Individuals	17744		
No. of deaths	1220		

Table 7: GLS Fixed-Effect Estimates of the Determinants of Various Health Status Measures

	Self-Assessed Health		**Difficulties with Chores**		**Mental Health**		**BMI**		**Serious Illnesses**	
	Coeff.	**|*t*|**	**Coeff.**	**|*t*|**	**Coeff.**	**|*t*|**	**Coeff.**	**|*t*|**	**Coeff.**	**|*t*|**
Age	0.002	0.46	0.050	7.19	0.017	2.43	0.007	0.84	0.120	51.11
No. of times married	0.031	0.45	-0.083	0.63	0.131	0.96	-0.114	0.69	-0.028	0.62
No. of times divorced	-0.048	0.61	-0.252	1.69	-0.071	0.46	0.444	2.36	-0.042	0.83
Regular exercise at *t*	0.120	9.32	-0.271	11.03	0.116	4.55	-0.152	4.92	-0.032	3.78
Regular exercise at *t*-3	-0.015	1.20	0.075	3.07	0.006	0.25	0.022	0.72	0.002	0.21
No. of drinks at *t*	0.001	0.18	-0.003	0.97	-0.004	1.46	0.008	2.25	-0.004	4.09
No. of drinks at *t*-3	-0.002	1.15	0.004	0.92	-0.005	1.17	-0.005	0.96	0.002	1.43
Smoker at *t*	0.094	3.00	-0.319	5.34	0.113	1.83	-0.565	7.52	-0.179	8.83
Smoker at *t*-3	-0.018	0.61	-0.117	2.04	0.041	0.69	0.024	0.33	-0.018	0.94
Spouse’s health	0.036	4.57	-0.032	2.10	0.105	6.70	0.012	0.58	-0.013	2.55
No. of children	-0.026	1.91	0.049	1.86	0.011	0.42	0.050	1.49	0.006	0.73
No. of siblings	-0.004	0.89	0.015	1.78	-0.002	0.19	0.008	0.70	0.002	0.80
Log of household income at *t*	0.001	0.11	0.016	1.60	0.018	1.73	0.028	2.19	0.001	0.22
Log of household income at *t*-3	-0.002	0.45	-0.008	0.97	0.007	0.74	0.001	0.10	-0.002	0.73
Years of tenure in longest reported job	0.003	0.76	0.011	1.56	0.001	0.19	-0.009	1.03	-0.005	2.18
Total years of work	0.001	0.18	-0.012	1.12	0.016	1.35	0.091	6.50	-0.011	2.98
Managerial Oper.	0.076	2.05	-0.250	3.56	0.046	0.63	-0.029	0.32	-0.009	0.38
Profess. Operator	0.132	3.41	-0.250	3.39	0.203	2.65	-0.022	0.23	-0.059	2.34
Sales	0.182	4.69	-0.362	4.91	0.194	2.53	-0.004	0.04	-0.045	1.78
Clerical/admin	0.113	3.31	-0.200	3.05	0.072	1.07	-0.156	1.90	-0.072	3.23

Services: cleaning	0.081	0.98	-0.286	1.82	0.337	2.07	-0.169	0.85	-0.002	0.04
Services: security	0.264	2.82	-0.502	2.81	0.149	0.81	-0.156	0.69	0.058	0.96
Services: food prep.	0.073	0.96	-0.207	1.44	-0.014	0.09	-0.433	2.38	-0.051	1.03
Health services	0.090	1.16	-0.391	2.63	-0.007	0.05	-0.180	0.96	-0.024	0.47
Personal services	0.087	1.90	-0.133	1.52	0.171	1.89	-0.092	0.83	-0.072	2.41
Farming/fishing/forest	0.212	2.96	-0.323	2.36	0.073	0.52	-0.169	0.98	-0.033	0.71
Mechanics/repair	0.072	0.91	-0.202	1.33	0.034	0.21	-0.017	0.09	-0.028	0.55
Construction	0.072	1.05	-0.274	2.09	0.260	1.91	-0.156	0.95	-0.116	2.59
Precision production	0.143	1.94	-0.580	4.12	0.226	1.54	-0.252	1.42	-0.153	3.20
Machine operator	0.034	0.60	-0.400	3.75	0.130	1.18	-0.117	0.87	-0.114	3.16
Transport operator	0.215	3.95	-0.363	3.47	0.189	1.74	-0.036	0.27	-0.071	1.98
Other operators	0.146	2.04	-0.564	4.15	0.244	1.73	0.023	0.13	-0.137	2.96
Constant	1.965	9.63	-0.696	1.79	-1.843	4.56	24.003	48.93	-5.115	38.63
Fraction of residual variance in fixed effects	0.708		0.771		0.6255		0.928		0.919	
R2 of observables	0.123		0.030		0.072		0.016		0.264	
Number of observations	25273		25273		25273		25273		25273	
Number of individuals	9134		9134		9134		9134		9134	
Test for fixed-effects	YES		YES		YES		YES		YES	

Table 8: Probit Model of Mortality with Health Measures

	Coeff.	**|*t*|**	**ME**
Age	0.036	8.59	0.001196
Self-assessed health at *t*	-0.216	7.31	-0.00713
Self-assessed health at *t*-1	-0.068	2.25	-0.00224
Self-assessed health at *t*-2	0.0001	0.01	1.29E-05
Self-assessed health at *t*-3	-0.013	0.46	-0.00044
Difficulties with tasks at *t*	0.039	3.18	0.001282
Difficulties with tasks at *t*-1	-0.037	2.69	-0.00123
Difficulties with tasks at *t*-2	-0.008	0.55	-0.00026
Difficulties with tasks at *t*-3	0.019	1.35	0.000637
Mental health at *t*	-0.012	0.82	-0.00037
Mental health at *t*-1	0.018	1.24	0.00059
Mental health at *t*-2	0.023	1.63	0.000766
Mental health at *t*-3	0.037	2.47	0.001216
BMI at *t*	-0.078	7.85	-0.00258
BMI at *t*-1	0.006	0.46	0.000185
BMI at *t*-2	0.045	3.78	0.001485
BMI at *t*-3	0.006	0.59	0.000187
Physical condition at *t*	0.247	6.76	0.008169
Physical condition at *t*-1	-0.055	1.05	-0.00183
Physical condition at *t*-2	-0.033	0.57	-0.00109
Physical condition at *t*-3	-0.014	0.30	-0.00045
Regular exercise at *t*	-0.222	4.19	-0.00728
Regular exercise at *t*-1	-0.053	1.05	-0.00174
Regular exercise at *t*-2	0.101	1.96	0.003441
Regular exercise at *t*-3	0.015	0.25	0.000509
Smoker at *t*	0.052	0.57	0.001771
Smoker at *t*-1	0.017	0.17	0.000576
Smoker at *t*-2	-0.087	0.86	-0.00273
Smoker at *t*-3	0.275	3.20	0.010741
Log household income at *t*	0.006	0.32	0.000206
Log household income at *t*-1	-0.017	0.94	-0.00055
Log household income at *t*-2	-0.014	0.88	-0.00046
Log household income at *t*-3	0.001	0.08	0.000041
Total wealth at *t*	-6.26E-08	0.75	-2.07E-09
Total wealth at *t*-1	1.03E-08	0.13	3.39E-10
Total wealth at *t*-2	-1.35E-07	1.18	-4.47E-09

Total wealth at *t*-3	9.88E-08	1.09	3.27E-09
No. of drinks at *t*	-0.005	1.19	-0.00018
No. of drinks at *t*-1	0.014	3.59	0.000447
No. of drinks at *t*-2	0.004	0.73	0.000147
No. of drinks at *t*-3	-0.016	2.18	-0.00052
Constant	-3.224	8.37	
Log L	-2158.17		
Pseudo R2	0.184		
Individuals	17540		
No. of deaths	1220		