Disentangling the relationship between health and income.*

Andrew M. Jones [†] and John Wildman [‡]

Abstract

The nature of the relationship between health and income is still not clearly defined. It is believed that personal income affects health but this may not be the end of the story. Income may also play its part through issues of relativity and deprivation. It may even be possible that these latter indirect effects are more important determinants of health than individual income. This paper investigates these issues by focusing on the relationship between income and health. The models are estimated using parametric and semiparametric panel data frameworks in order to ameliorate problems of misspecification and unobservable heterogeneity. The results demonstrate strong evidence for income affecting health but limited evidence for relative deprivation.

JEL: I12, C14, C23
Keywords: Health production, panel data, semiparametric methods, absolute income, relative deprivation.

*The BHPS data were supplied by the: Economic and Social Research Council Research Centre on Micro-Social Change, British Household Panel Survey. Colchester, Essex: The Data Archive, 22 December 1998. SN: 3953. Neither the original collectors of the data nor the Archive bear any responsibility for the analysis or interpretations presented here. This paper derives from the project "The dynamics of income, health and inequality over the lifecycle" (known as the ECuity III Project), which is funded in part by the European Community's Quality of Life and Management of Living Resources programme (contract QLK6-CT-2002-02297). The authors acknowledge the comments of participants at iHEA Fourth World Congress and York Seminars in Health Econometrics. All remaining errors are ours.

[†]Department of Economics and Related Studies, University of York, Heslington, York YO10 5DD; email: amj1@york.ac.uk tel: +44 1904 433766, fax: +44 1904 433759

[‡]Economics, University of Newcastle, Newcastle upon Tyne, NE1 7RU, UK; email:j.r.wildman@ncl.ac.uk, tel: +44 191 222 8645, fax: +44 191 222 6551

1 Introduction

The nature of the relationship between health and income is still ill-defined. Few researchers doubt that income is an important determinant of health but there are many schools of thought on how exactly the relationship functions. It seems reasonable to believe that own income is an important determinant, with own income being directly linked to individual health. However, there are a number of indirect mechanisms that may or may not be important determinants of health (Deaton, 2003). These indirect links may be due to income inequality, relative deprivation, social capital or other equally complicated pathways which are difficult to define and to test for. Evidence of ill-health being generated by hierachy is common in epidemiology and biology (Marmot et al., 1978; Sapolski, 1993).

Economics has a long tradition of incorporating relative income and status into models of individual utility maximisation. Duesenberry (1949) hypothesised that the individual want for self-esteem makes them emulate the consumption patterns of those who are on a higher socioeconomic rung of the ladder. Frank (1985a, 1985b, 1997) has also considered the impact of positional goods and subjective well-being. Frank (1997) contrasts the psychological approach to well-being with that of economists. Economists classically speak of individual utility, with utility being a function of present and future consumption, leisure etc. The individual is assumed to choose between alternatives to maximise utility. Within the utility function variables relating to status or relative position are not included. Given that individuals are driven by the need for status such assumptions are unrealistic (Frank, 1985a). Empirical work in economics has also started to investigate issues of interdependent utility. Clark and Oswald (1996) include mean income levels in models of happiness to control for comparisons. Clark (2003) has recently investigated the relationship between well-being and income inequality, finding that income inequality increased well-being. Blanchflower and Oswald

(2003, 2004) using US data find that inequality does reduce happiness.

Health production functions are analogous to this argument; individual health has been specified as a function of individual characteristics (Grossman, 1972) with little or no concern for relative position. If seeking relative status causes stress, increased working hours and general unhappiness, then these considerations need to be accounted for. This point is not unrelated to utility, individuals who describe themselves as happy (have higher utility) suffer from less psychosomatic illnesses; are less likely to be absent from work; less likely to die prematurely; less likely to attempt suicide; less likely to seek psychological counselling (Frank, 1985b; Clark and Oswald, 1996. If health is vital for a flourishing life (Culyer, 1993) then understanding the pathways between relative position and health is vital.

Even the direct relationship between health and income is poorly understood. It is widely acknowledged that income determines health in a non-linear fashion, with the standard assumption of diminishing marginal returns to increasing income. However for model specification, this may be a strong assumption and lead to biased coefficients in the estimated model. It is also possible that there are effects of reverse causation, with health affecting individual income as well as income affecting health. Individuals who are healthier may have higher incomes. It may be that individuals with low discount rates invest in health and invest in earnings potential. Such endogeneity is controlled for by parameterising the individual effect.

This paper uses a range of panel data estimation techniques to investigate the impact of both income and relative deprivation on health. The results demonstrate that income is an important determinant of health for both men and women, with limited evidence for relative deprivation.

2 Examining the relationship between health and income.

Traditionally individuals are assumed to produce health using their own inputs, with no interdependence between individuals (Grossman, 1972).[1] Although this paper will not consider a full health production model one can analogously apply the arguments to the determinants of health. Should the determinants of health include individual variables or should some form of interdependence be taken into account?

Income can directly affect health; it seems likely that my level of income will impact on my level of health. An individual with higher income can partake in a healthier lifestyle or purchase goods and services to alleviate or prevent the onset of ill health. It may be that lower income leads to lower health due to financial commitments or stress.

However, these direct effects of income may mask more complicated indirect effects. While income may act as something that allows inclusion in society, a lack of income may lead to exclusion. Such ideas are echoed in the writings of Townsend (1979) and his interpretation of relative poverty. These feelings may encompass deprivation effects or the impact of income inequality on individual health. If income inequality is affecting individual health then it can be demonstrated that any policy attempting to reduce health inequality must focus on narrowing income inequality by raising the incomes of those in the poorest groups of society (Wildman 2003). The idea of trickle down will not work if income inequality is important.

Sociologists have long pondered the issues of interdependent factors in the determinants of health. Townsend's relative poverty work demonstrates the role that exclusion can play on individuals, this theme was picked up by

[1] This seems remarkable in itself given that health is considered to be a unique area of study due to issues such as externalities. The health of Janet affects the health of John, due to infection, vaccination behaviour etc.

Desai and Shah (1988) in the economics literature to try and quantify such relative inequality measures. More recently Wilkinson (1996) has suggested that income inequality has a larger impact on individual health than income itself. The pathways by which this effect could work are complex and unclear (Wildman and Jones, 2001). Income inequality may have some independent impact on health, or it may function through issues of control and stress. Societies that are less uneven are more likely to have strong support networks that allow recovery or ameliorate the impact of stress and lack of control on health. Such polarization issues are reflected in the work of economists who have tried to identify whether there is a 'disappearing middle' (Esteban and Ray, 1994).

This paper presents a framework to empirically investigate these issues. The purpose is twofold, one to further investigate the relationship between health and income and secondly to test whether relative deprivation is having a negative impact on individual self reported health. Relative deprivation is acting here as a proxy for the factors which may cause relative status to impact upon individual health.

Any paper investigating these issues needs to cope with a number of specification issues. Firstly, it may be the case that individuals with high incomes have good health due to some unobserved characteristic. Such a characteristic could be discount rates, an individual may have low discount rates and so invest in their future health and also invest in their earnings potential. Such a characteristic would lead to the observation that individuals with high income have good health. To counter this problem we may include variables to control for such unobservables, one obvious candidate in this case is education. An individual with high levels of education may have low discount rates because high levels of educational attainment indicate investment in human capital. While this may provide a useful control it may not fully control for the unobservable, or it may not have any impact in controlling for some other unobservable that is not correlated with education. We need to have some mechanism to control for such individual heterogeneity and in such cases panel data provides that control. Since we have repeated observations on the same individual we can control for any time invariant unobservable characteristics, so if we assume time discount rates are constant across time periods we can sweep them out of the model either by using the fixed or random effects panel data approach.

The second issue is the nonlinearity of the relationship between health and income. If the relationship between health and income is nonlinear but we specify a linear model then the extra nonlinearity acts as an omitted variable and if correlated with other variables will lead to biased estimates. This will be especially true of the coefficient attached to any deprivation variable which is likely to be correlated with income. For example, if the true nature of the relationship between health and income were a quadratic we would want to estimate a model including both income and income squared. If however we estimate a model using only income we have essentially omitted income squared. This term may cause three problems, firstly the coefficient on income will not reflect how income actually determines health, it may over- or under-estimate its impact. The second problem is that the correlation with the omitted variable and other included variables will lead to the coefficients on these other variables being biased. There are also problems with estimating standard errors; by omitting a variable the covariance between the omitted variable and the included variables is included in the estimation procedure leading to smaller standard errors on included variables and a higher probability of making Type I errors.

This argument is more general, however many nonlinear terms we include in our regression model there is still the possibility that an extra nonlinear term has been omitted, leading to the problems described above. To overcome these issues we estimate a model using semiparametric methods. This allows us to leave the nature of the relationship between health and income to be unspecified and so overcome any specification issues. This gives us the best chance of obtaining an unbiased coefficient on the deprivation variable and also allows us to observe the actual relationship between health and income. Further, we compare the results from the semiparametric estimation with results from models where income is specified in logs. This allows us to assess the validity of this functional form.

3 Methods

The models are estimated using fixed and random effects approaches. Random effects estimation controls for unobservable heterogeneity and leads to efficient estimators as long as the exogenous variables are uncorrelated with the error term. If correlation is present then the coefficients will be biased and inconsistent. To test for correlation we use a Hausman t-test (Wooldridge, 2002). A standard Hausman test is based on comparing all the estimated coefficients from fixed and random effects models. Often we are only interested in certain variables in the regression model in which case it makes sense to test these variables on an individual basis. The Hausman test in this case can be computed as $\hat{\beta}_{FE} - \hat{\beta}_{RE}/([se(\hat{\beta}_{FE}]^2 - [se(\hat{\beta}_{RE}]^2)^{1/2}$, where the standard errors are computed under the usual assumptions. In cases where the random effects model fails a Hausman test we use a Mundlak (Mundlak, 1978) formulation to control for the correlation between the exogenous variables and the error term. This involves parameterising the individual specific effect using the time series means of the regressors. The models were then retested. We also estimated the parametric models using a Hausman-Taylor estimator (Hausman and Taylor, 1981) which fits random effects panel data models where some of the covariates are correlated with the unobserved individual level random effect. It also allows us to estimate models with potentially endogenous time invariant variables.

3.1 Parametric models

Linear models are used to estimate the determinants of health model. The use of semiparametric methods will allow for a general functional form between health and income, which includes the logistic model and other non-linear estimators. We initially specify a parametric model of the form:

$$H_{it} = X_{it}'\beta + u_i + \varepsilon_{it} \tag{1}$$

where H_{it} is the health of individual i at time t, X_{it} represents a set of explanatory variables which may affect health, including income and a measure of relative deprivation. The βs are vectors of parameters to be estimated. The error term, ε_{it} is time and individual specific, is uncorrelated with X_{it} and u_i and across individuals and is assumed to be drawn from a normal distribution with mean zero and constant variance σ_ε^2. The u_is represent the individual specific and time invariant random error component, assumed to be drawn from a normal distribution with mean zero and a constant variance σ_u^2.

The Hausman-Taylor technique estimates a model of the form:

$$H_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + Z_{1i}\delta_1 + Z_{2i}\delta_2 + u_i + \varepsilon_{it} \tag{2}$$

where X_{1it} is a 1 x k_1 vector of observations on exogenous, time-varying variables assumed to be uncorrelated with u_i and ε_{it}; X_{2it} is a 1 x k_2 vector of observations on endogenous, time-varying variables assumed to be (possibly) correlated with u_i and ε_{it}; Z_{1i} is a 1 x g_1 vector of observations on exogenous, time-invariant variables assumed to be uncorrelated with u_i and ε_{it}; Z_{2i} is a 1 x g_2 vector of observations on endogenous, time-invariant variables assumed to be (possibly) correlated with u_i and ε_{it}; u_i is the unobserved random effect and ε_{it} is the idiosyncratic error term.

3.2 Semiparametric models

Parametric estimation may lead to biased coefficients if the parameterisation of the explanatory variables is incorrect. Extra nonlinearity in the relationship between health and income may be passed on to other estimated coefficients. Semiparametric partially linear estimation (Robinson, 1988) allows us to overcome this problem by allowing an unspecified relationship between health and income [2].

We specify a partially linear model of the form:

$$H_{it} = g(y_{it}) + X_{it}'\beta + u_i + \varepsilon_{it} \tag{3}$$

In this model the matrix X_{it} contains all the explanatory variables except income and $g(y_{it})$ is the unknown function of income. Using the methods of Robinson (1988) we transform the model by taking expectations of (3) and then subtracting the expectations from (3). This yields:

$$H_{it} - E[H_{it}|y_{it}] = (X_{it} - E[X_{it}|y_{it}])\beta + \epsilon_{it} \tag{4}$$

To implement the estimator the expected values are replaced by their nonparametric estimates. We use locally weighted regression to estimate these nonparametric functions (Härdle, 1990). Estimation of (3) leads to $\sqrt{N}$ consistent estimators of the beta coefficients (Robinson, 1988). While the asymptotic properties of these estimates are well founded, in finite samples the estimators may perform poorly. The functional relationship between health and income can be retrieved by a nonparametric regression of $(H - X'\beta)$ on y_{it}.

[2]Such methods could not be estimated for Hausman-Taylor type models.

4 Data

4.1 The BHPS.

The models are estimated using 11 waves of data from the British Household Panel Survey (BHPS). The BHPS was chosen because it is a recent panel data set with good quality income and health variables.

The BHPS is a repeated panel: respondents are questioned each year, containing socio-demographic, income and health variables. The BHPS is a longitudinal survey of private households in Great Britain (England, Wales and Scotland) and was designed as an annual survey of each adult member (16+) of a nationally representative sample of over 5,000 households, giving 9,902 individual interviews. The initial sample, collected in 1991.

The longitudinal nature of the BHPS requires that individuals are re-interviewed each year. All individuals who were interviewed at wave one were re-interviewed (unless they dropped out, e.g. because of death or lack of cooperation) and these are known as the Original Sample Members (OSM's). The sample for other waves consists of all adults in all households containing as least one member who was resident in a household interviewed at wave one. It is hoped that this will keep the sample broadly representative of the population of Britain. Information at both the household and individual level was collected, covering questions on neighbourhood, income, employment, health and caring, demographics, and values and opinions. [3]

[3] Although attrition may be a problem with panel data estimation, it was found by Contoyannis et al. (2004) that attrition, while a problem in the BHPS, was not affecting the estimation results for similar models of determinants of health.

4.2 The Dependent Variable

We use two dependent variables, a continuous measure of psychological well-being derived from the General Health Questionnaire (GHQ)[4] and a dichotomised measure of self assessed health (SAH).

The self completion questionnaire component of the BHPS includes a reduced version of the (GHQ) (Goldberg and Williams, 1988). The GHQ was developed as a screening instrument for psychiatric illness but is often used as an indicator of psychological well-being. There are 12 individual elements in the shortened GHQ covering concentration, sleep loss due to worry, perception of role, capability in decision making, whether constantly under strain, perception of problems in overcoming difficulties, enjoyment of day-to-day activities, ability to face problems, loss of confidence, self-worth, general happiness and whether suffering depression or unhappiness. Respondents indicate on a four point scale (ranging from 0 to 3, 0 being the best score) how they have recently felt when responding to each item. For the dependent variable the Likert scale is used (Likert, 1952). The Likert scale obtains an overall score by summing the individual components of the GHQ, which ranges from 0 to 36 and is increasing in good health.

The GHQ variable is a useful measure of health for determining the effect of income related variables because it measures psychological well-being and mental well-being which 'concentrates on broader components of psychiatric morbidity (particularly anxiety and depression). It was not intended

[4]The GHQ score has recently been interpreted as a measure of happiness in models of individual utility (Clark and Oswald, 1996; Clark, 2003). Our interpretation is different, we interpret GHQ as a measure of health in a determinants of health model, rather than an indication of happiness or utility. Although the model specification is similar there are subtle differences in the interpretation of the results. We can compare models of health determinants between measures of well-being, which detect psychological health (GHQ scores) and measures of self-reported health which are more highly correlated with mortality and physical health.

to be used for detection of functional psychoses (schizophrenia or psychotic depression) although these conditions are in fact detected' (Bowling, 1991) Detecting broad components of psychiatric morbidity is important if income and deprivation are affecting individuals via stress or psychosocial mechanisms (Wilkinson, 1996).

SAH is ascertained by asking the question 'Please think back over the last 12 months about how your health has been. Compared to people of your own age, would you say that your health has on the whole been excellent/good/fair/poor/very poor?' We dichotomise our variable by assigning a value of one to those in good health or better and zero otherwise. This allows the use of linear probability models in the estimation procedure. SAH should be interpreted as perceived health status relative to the individual's concept for the 'norm' for their age group. It is a widely used measure of health status (eg Adams et al., 2003, van Doorslaer et al., 1997, Ettner, 1996) and has been shown to predict inequalities in mortality (see e.g., van Doorslaer and Gerdtham (2003)).[5]

4.3 The Independent Variables

4.3.1 Income Variable

The income variable used in the empirical models to measure absolute income is annual household income, deflated and equivalised using the McClements equivalence scales. The income variable FIHHYR is included in the BHPS as a derived variable, which gives annual total household income in the reference year. The reference year is defined as the twelve months prior to the start of the interview period (the 1 September of the year in question).

In the semiparametric models no functional form is imposed on income. In the parametric models the log of income (LNY) is used to allow a non-

[5]For an in-depth investigation into SAH and its dynamics see Contoyannis et al., 2003.

linear relationship between health and income, as suggested by previous work (Ettner, 1996; Ecob and Davey Smith, 1999). [6]

4.4 The relative deprivation measure.

When constructing any measure of deprivation it is important to consider the reference group problem. With whom do individuals make comparisons? In the literature to date, the issue of reference groups has been largely unresearched. Wilkinson (1996) uses country level inequality measures, implying that the whole country is the level of the reference group. Kaplan et al. (1996) and Kennedy et al. (1996) use US state level inequality measures, suggesting that in the US states are the correct reference group. Soobader and LeClere (1999) investigated the impact of income inequality at a variety of aggregation levels and find that the impact on income inequality is higher at higher levels of aggregation. This is consistent with the hypothesis that at low levels of aggregation there will be too little heterogeneity between individuals to result in inequality being a factor affecting health (Wilkinson 1996). Clark (2003) uses a reference group based on individuals of the same sex, region and year. Blanchflower and Oswald (2003) use a measure of relative income based on individual income divided by state income. Lindley and Lorgelly (2003) use reference groups based on standard regions in the UK. Although, considering reference groups as purely geographic seems to neglect how individuals may make comparisons. Runciman (1966) suggests two possible reference groups: individuals in the same social situation, e.g. comparison within social class or neighbourhood, and comparison with

[6] We use household income as our measure of income even though this may not be representative of the true income an individual receives, it is dependent on the household allocation of resources. The use of individual income is also problematic, however, since individuals in high income houses may receive no income, if they are not an earner, and so be recorded as having an income of zero. Also individual income is likely to be measured with greater error than household income and so provide less robust results.

individuals you wish to emulate.

Deaton (2001) suggests that the most sensible reference groups for the US are either the whole population or those that live in an immediate geographical location. Our data set does not provide enough observations to consider small geographic areas, so we consider the whole country as the reference group, each individual compares him or herself with every other individual.

There may also be differences between conscious and subconscious comparison and deprivation. It may be the case that individuals make comparisons at both levels. By explicitly formulating a reference group for individuals, in a regression model, based on locality or similarity in sociodemographic characteristics one, is boxing the individual into conscious comparison and may be neglecting subconscious issues. By using the whole population as the reference group we are at worst watering down the comparison (individuals may feel the comparisons more when individuals are closer to them).

The deprivation measure constructed is based on the deprivation measure described in Hey and Lambert (1980). Hey and Lambert (1980) define D(m;y), the level of deprivation felt by individual i with income m with respect to income y, as,

$$D(m;y) = \begin{cases} y - m & : \quad if \quad m < y \\ 0 & : \quad if \quad m \geq y \end{cases} \tag{5}$$

The individual feels more deprived as the number of individuals in society with income y increases. An overall measure of deprivation for the individual is given by weighting the measure by the proportion of society with income y.

The deprivation for an individual with income y is calculated from the formula (Chakravarty, 1990):

$$d_y(F) = \mu[1 - F_1(y)] - y[1 - F(y)], \tag{6}$$

where μ is mean income, $F_1(y)$ is the cumulative proportion of total income at the income y and F(y) is the cumulative proportion of the population up to the individual with income y (where the population is ranked by income).

A deprivation measure for the whole sample is created, which we call deprivation for the population (RDEP1). For this measure deprivation is 0 for the individual with the highest income and positive otherwise.[7]

4.5 Other Independent Variables

To allow for a flexible relationship between the SAH and age we categorise age according to 6 age groups.

Variables for household size (HHSIZE) and number of children (NKIDS) are included. Regional dummies are included to control for any regional effect. Indicators of marital status are included in the model (WIDOWED (widowed), DIVSEP (divorced or separated) and NVRMAR (never married) are included in the models, married is always the excluded category).

Job status is included in the model, the categories are SELF (self-employed), UNEMP (unemployed), RETIRED (retired), DISAB (disabled), FAMCARE, MATLEAVE and OTHER (other groups); the category 'employed by others', was always excluded. Education is measured as the highest qualification obtained over the sample period in descending order of achievement (Degree, alevel/hnd, olevel/cse and no qualifications).

The full set of explanatory variables and their sample means are summarised in TABLE 1.

5 Results

The results presented are from unbalanced panels.

[7]Summing the relative deprivation scores of each individual gives the Gini coefficient, a well known measure of inequality.

5.1 Women

For SAH (Table 2), the fixed effects results (columns 1 and 6) indicate that income has a significant positive impact for women, with an increase in income increasing the probability of reporting good health by 1%. The inclusion of the relative deprivation variable does not alter these results. When relative deprivation is included it is positive and insignificant, suggesting no impact on individual health.

Being unemployed, retired a family carer, student or disabled has a negative and significant impact on individual health. Women who are self employed or on maternity leave do not have significantly different health to working women. The age variables are all positive and significant, indicating that health, compared to individuals of the same age, is increasing through the age groups. The year dummies indicate that generally, compared to 1991 the reporting of SAH is falling through time, the only outlier here is the coefficient attached to the 1999 dummy. This was wave 9 of the BHPS where the question for SAH was altered as part of the SF-36 questionnaire. [8]

Specifying the random effects model was a difficult process. Initially including the population means for the job status variables only (columns 2 and7) meant that the coefficients on income and on relative deprivation did not both pass the Hausman test. When relative deprivation was included in the model (column 7) the income variable passes but the deprivation variable did not.

The next step was to include the time series average of the income variable as a control, for parsimony this was the preferred model not including relative deprivation (column 4). In the model without deprivation the income variable passed the Hausman test. In the models including relative

[8] Including a dummy for the ninth wave seems to be the best method for controlling for this anomaly (Hernandez-Quevedo et al.,2004).

deprivation, income passed but the deprivation variable failed. So a further specification was tried with the time series average of relative deprivation also included (column 9). In this case the income and the relative deprivation variables pass the Hausman test at the 10% level. This was our preferred model with relative deprivation included. The models which parameterise the individual effect also control for potential endogeneity within the model that may stem from the unobserved individual effect.

In column (2) we see that income has a positive and significant impact on SAH. The magnitude of the variable is almost double that of the fixed effects model. Those classed as unemployed, retired, family carers, student and disabled again reported significantly worse health compared to women who are employed, although the coefficient on the disabled variable failed the Hausman test. When relative deprivation is included the impact of income is unaffected in magnitude, although it is no longer significant. This may be the result of collinearity between the variables. Relative deprivation itself is negative but not significant. The random effects model allows us to estimate the impact of education on health; in this case we see a clear education gradient. More education leads to better health, having a degree increases of the probability of reporting good health by about 13% compared to having no qualtification. Also we see that women who are nonwhite report worse health than whites, with nonwhite around 7%. The other variables in the model remained largely unchanged, with the same subset of variables significant and the same subset passing the Hausman test. Interestingly, the time series average of relative deprivation included as a control in this model was negative and highly significant, suggesting that there may be some long run mechanism through which relative deprivation is affecting health.

The results for the GHQ variable are presented in Table 3. In the fixed effects models (columns 1 and 6) income is positive and significant. When relative deprivation is included the coefficient on income is larger and still significant, although relative deprivation itself is positive and insignificant. Being unemployed, retired, on maternity leave, a family carer or disabled significantly reduces your health compared to women who work. Being unemployed leads to a fall in the GHQ score of more than one and a half points compared to a women who works. Women who are widowed or divorced have significantly worse health then women who are married, although women who are not married report significantly better health. It also seems to be the case that women in the older age groups are more likely to report significantly better psychological well-being than younger women. The time dummies again indicate that, over time, worse health is being reported.

For the random effects model we followed the same procedure for specification as outlined above. The standard random effects model (allowing for no Mundlak-style correction) results in estimates on income that pass the Hausman test, although many of the job status variables fail (columns 2 and 7). This is true for models including and excluding relative deprivation. When relative deprivation is included, the coefficient on relative deprivation fails the Hausman test. Including the time series average of deprivation in the regression means that relative deprivation passes the test (column 9). In all models income is still positive and significant, while relative deprivation is positive and insignificant in the model it passes. Although, like SAH, the time series average of relative deprivation is negative and significant. The other variables in the model generally retain their sign and significance. The education gradient can again be observed, although for GHQ the coefficient attached to ocse is greater than that attached to ahnd. For nonwhites there is not a significant difference in reported GHQ scores.

The results from the Hausman-Taylor estimation are generally supportive of the random effects models (Table 4). For SAH income is positive and significant both when relative deprivation is included and excluded, with estimates similar those from the random effects models reported above. Relative deprivation is positive and insignificant. The same job status variables are negative and significant and the estimates for nonwhite are also negative and significant. The estimates for the time invariant endogenous education variables are interesting, the gradient is still observed, with coefficients of smaller magnitudes to the results above, but the coefficient on degree is now insignificant. This suggests that increasing education does not have increasing returns on health.

For the GHQ the results follow a similar pattern. Income is always positive and significant, with coefficients of similar magnitude to the random effects models. The other socioeconomic variables follow the same patterns as the random effects estimates. The results for education are slightly different. Women with the top (DEGREE) and middle (OCSE) levels of education do not have significantly different reported GHQ scores compared to women with no qualifications. Women with the middle level of qualifications (AHND) have significantly better GHQ scores.

5.2 Men

The same estimation procedure outlined above was used for men as well. For SAH (Table 5) the fixed effects models show that income has a positive and significant impact on individual health (columns 1 and 6), the magnitude of the results are similar to those for women, with income increasing the probability of reporting good health by 1%. When relative deprivation is included in the model the size of the coefficient increases, although deprivation itself is positive and insignificant. Men who report a job status apart from self employed or the Other category, report significantly worse health then employed men. Men who have never married report significantly worse SAH then married men. The time dummies are generally negative and significant, apart from the results for 1999.

The random effects models generally support the fixed effects results, although when relative deprivation is included the coefficient on income is insignificant (in specifications where both income and relative deprivation pass the Hausman test (column 7)). The time series average of relative deprivation was negative and significant when included in the model, indicating some possible effect of relative deprivation.

The education gradient is clearly visible, men with higher qualifications report better health. The coefficient attached to nonwhite is also negative and significant.

For the GHQ (Table 6), income is not significant in the fixed effects models and is only significant in one of the random effects models, although in this case it did not pass a Hausman test. Apart from the self employed, all other job status categories reported worse GHQ scores than men who were employed. Relative deprivation is negative and insignificant in all the models, indicating no deprivation effects. Although again, the time series mean of relative deprivation is negative and significant. There is no significant difference between whites and nonwhites, also the education gradient has disappeared, with individuals with the lowest levels of qualifications having significantly better health than individuals with no qualifications in some models and individuals with degrees reporting better GHQs in other models. The time dummies show that GHQ levels are falling over time.

The Hausman-Taylor results (Table 7) are slightly different to the results above. For SAH income is positive and significant whether relative deprivation is included or not. The magnitudes of the estimated coefficients are similar to the results from the fixed and random effects models. The relative deprivation variable is positive and insignificant. The other results are similar to the random effects results, although the coefficient estimated for nonwhites is insignificant. The education gradient is still present in terms of the size of coefficients, although all the estimated education coefficients are smaller than those from the random and fixed effects models, the coefficient attached to degree is now insignificant.

For GHQ, there is no impact for either income or relative deprivation. The other estimated coefficients are largely the same as the for the random effects models. The education variable is different however and we observe that there is only a significant impact for the middle level of qualification (ahnd).

5.3 Semiparametric Results

The semiparametric results for women largely support the parametric results. For SAH the coefficient on deprivation is negative but insignificant.[9] The impact for the other included variables remains fairly unchanged. For the GHQ deprivation is again negative but still insignificant, the results for the other variables are similar, although the magnitudes are generally smaller. Figures 1 and 2 demonstrate that the relationship between SAH and income is not linear, SAH initially decreases as income increases, SAH then rises quickly as income increases further, until the returns to increasing income begin to flatten out. For the GHQ, the relationship with income is more linear and flatter, although there does seem to be a slight decrease in psychological well-being initially as income increases followed by small improvements. [10]

For SAH for men we find that the results are similar to the parametric models. Relative deprivation is positive but insignificant. The results for the other covariates are generally the same, although the size of the coefficients are generally smaller. For GHQ we find a different result to the parametric models. The impact of relative deprivation is negative and significant, this is the first indication that relative deprivation makes an impact on individual

[9] These results were estimated with different bandwidths but the results were largely unchanged.

[10] The smoothed relationships between health and income have been overlaid with the density of income to demonstrate where the majority of the observations are occuring.

health, in this case psychological well-being. The coefficient has changed sign and become more significant after allowing for the flexible relationship between health and income. The shape of the relationships between health and income are similar to those for women (Figures 3 and 4).

6 Discussion

The results presented provide further insights into the relationship between income and health. For women, for SAH and GHQ there is a significant relationship between health and the log of income, with income improving health. These results are supported by the Hausman-Taylor estimates and by the semiparametric results. The semiparametric results indicate that the relationship between GHQ and income is linear and fairly flat. The relationship between SAH and income is much more nonlinear, this is reflected in the fact that many of the semiparametric estimates differ from their parametric counterparts, the relative deprivation variable is more negative in the semiparametric models. The coefficients on the education variables are smaller in the semiparametric models, in fact they are closer to the estimates from the Hausman-Taylor models. There are no deprivation effects in the any of the models estimated.

For women both mental and physical health is better for those who worked compared to other job status categories, except those on maternity leave for SAH and students for GHQ. Those who are unemployed, retired, family carers and registered disabled report the worse health for both SAH and the GHQ. There is evidence that women who are divorced report worse SAH compared to married women. Women who are divorced/separated and widowed report worse psychological well-being than married women, although women who have never married report better health for the GHQ. Interestingly, the larger number of children leads to higher reported SAH scores but has no effect for psychological well-being (GHQ). It is also clear that nonwhite women suffer from lower SAH than white women.

The education gradient is clear in all of the estimated models. The Hausman-Taylor results however, suggest that the gradient may be not so clear cut and may be due to the correlation between the endogenous education variables and the individual effect. In these latter models not all of the educational variables are significant.

For men the relationship between health and income is more mixed. While there is a clear significant relationship between health and income when health is measured as SAH (although this disappears when relative deprivation is included in the random effects model), no such relationship exists when using the GHQ. There is clear evidence that men who have a job category other than employed or self-employed suffer from worse health. There do not seem to be many effects of marital status for SAH, although there is some evidence that single men report worse health than their married counterparts. For GHQ there is evidence that all marital status catagories lead to worse health compared to married men (although the Hausman-Taylor results suggest there may be no impact for non-married men). Nonwhite men also suffer from worse SAH than married men.

The semiparametric results largely support the other models for SAH. For GHQ there are some interesting results, the most important being the significant negative impact of relative deprivation. The results suggest that the parametric specification of income and GHQ is incorrect and that when we overcome the specification bias we find that relative deprivation does play an important role in determining psychological well-being for men. These results demonstrate the importance of specifying the correct functional form of health and income.

6.1 Conclusion

This paper has presented a further investigation into the relationship between two measure of health and income. By exploring a range of panel data methods we have been able to make a large contribution to the literature in this field. We find that the significant relationship between SAH and income for men and women is clearly established, as is that between GHQ and income for women (for men no such significant relationship seems to exist). This supports the work of Ettner (1996) who finds a nonlinear relationship between health and income. We have also investigated the impact of relative deprivation on health. We find that there are no effects for women using either health measure. In the parametric models there is also no impact for men, which support the review of Lynch et al. (2004) who conclude that there is little evidence in support of income inequality affecting health. However, the use of robust semiparametric techniques has demonstrated that relative deprivation has a large and significant negative impact on the psychological well-being (GHQ score) of men. As we argued previously, it is more likely that deprivation effects would be detected in measures of mental health/psychological well-being, rather than physical health measures. This is a clear indication of that phenomenon. Although, these results contradict those of Clark (2003) who using BHPS data finds that inequality improves well-being. The differences may come from Clark's use of individuals who work and from the pooling of the data. When the models are re-estimated exploiting the panel nature of the data Clark (2003) finds no relationship between inequality and the GHQ scores. Blanchflower and Oswald (2003) however find that inequality does reduce happiness in the US.

We also investigated the impact of other socioeconomic variables on our health measures. We find that job status and marital status are important determinants of health. We also find that the relationship between health and education is important, although, as the Hausman-Taylor results demonstrate, we have to allow for the possible endogeneity of the education variables due to unobservable heterogeneity. The observed education gradient in standard random effects models may not be a true reflection of the impact education has on health.

Table 1: Summary Statistics

Variable	Women n=61402 T=11	Women n=62925 T=11	Men n=51908 T=11	Men n=53084 T=11
GHQ	24.164		25.514	
SAH good		0.689		0.735
Log income	9.532	9.526	9.633	9.627
Relative deprivation	-1.238	-1.231	-1.356	-1.349
No. kids	0.612	0.608	0.562	0.560
Household size	2.772	2.764	2.870	2.864
Self employed	0.036	0.036	0.111	0.111
Unemployed	0.027	0.027	0.060	0.060
Retired	0.210	0.215	0.183	0.187
Maternity leave	0.023	0.023	0.000	0.000
Family carer	0.145	0.146	0.007	0.007
Student	0.038	0.037	0.039	0.038
Disabled	0.031	0.032	0.044	0.044
Other	0.004	0.004	0.005	0.005
Widowed	0.117	0.121	0.036	0.037
Divorced/separated	0.087	0.087	0.054	0.054
Never married	0.157	0.156	0.208	0.208
age3039	0.211	0.209	0.212	0.210
age4049	0.179	0.177	0.183	0.182
age5059	0.143	0.143	0.149	0.148
age6069	0.111	0.112	0.112	0.113
age70	0.142	0.147	0.113	0.117
South West	0.079	0.079	0.083	0.083
London	0.083	0.083	0.083	0.082
Midlands	0.151	0.151	0.157	0.157
North West	0.093	0.092	0.094	0.093
North	0.140	0.139	0.139	0.139
Scotland	0.143	0.144	0.134	0.134
Wales	0.103	0.105	0.102	0.103
Degree	0.108	0.107	0.138	0.137
Ahnd	0.207	0.205	0.276	0.274
Ocse	0.304	0.301	0.269	0.267
Nonwhite	0.029	0.031	0.030	0.032
y1992	0.077	0.078	0.078	0.078
y1993	0.074	0.074	0.075	0.074
y1994	0.074	0.074	0.076	0.075
y1995	0.073	0.072	0.074	0.073
y1996	0.075	0.075	0.076	0.076
y1997	0.088	0.089	0.089	0.089
y1998	0.088	0.087	0.086	0.086
y1999	0.125	0.126	0.123	0.124
y2000	0.124	0.123	0.122	0.122
y2001	0.120	0.120	0.120	0.120

Table 2: SAH fixed and random effects models: Women

	1	2	3	4	5	6	7	8	9	10
	FE	RE m1	H	REm2	H	FE rdep	RE m1	H	REm2	H
Log income	0.009**	0.016**	f	0.008**	p	0.013**	0.008	p	0.007	f
Relative deprivation						0.004	-0.008*	f	-0.001	f
No. kids	0.017**	0.015**	p	0.016**	p	0.017**	0.016**	p	0.017**	p
Household size	-0.005	-0.001	f	-0.001	f	-0.005*	-0.001	f	0.000	f
Self employed	0.011	0.011	p	0.011	p	0.011	0.011	p	0.011	p
Unemployed	-0.030**	-0.029**	p	-0.030**	p	-0.030**	-0.029**	f	-0.030**	p
Retired	-0.041**	-0.040**	p	-0.042**	p	-0.041**	-0.040**	p	-0.043**	p
Maternity leave	-0.009	-0.004	f	-0.005	f	-0.009	-0.004	f	-0.005	f
Family carer	-0.043**	-0.041**	p	-0.043**	p	-0.043**	-0.041**	p	-0.044**	p
Student	-0.034**	-0.027**	f	-0.031**	p	-0.034**	-0.027**	f	-0.031**	p
Disabled	-0.183**	-0.186**	f	-0.188**	f	-0.183**	-0.186**	f	-0.188**	f
Other	-0.033	-0.034	p	-0.035	p	-0.033	-0.034	p	-0.036	p
Widowed	-0.020	-0.026**	p	-0.025**	p	-0.020	-0.026**	p	-0.024**	p
Divorced/separated	0.001	-0.008	p	-0.006	p	0.001	-0.007	p	-0.006	p
Never married	-0.005	0.022**	f	0.022**	f	-0.006	0.022**	f	0.023**	f
age3039	0.033**	0.006		0.005		0.033**	0.006		0.004	
age4049	0.065**	0.018**		0.017**		0.065**	0.018**		0.016*	
age5059	0.085**	0.025**		0.023**		0.085**	0.025**		0.022**	
age6069	0.122**	0.039**		0.037**		0.121**	0.040**		0.037**	
age70	0.097**	-0.011		-0.012		0.097**	-0.010		-0.012	
South West	0.030	-0.003		-0.001		0.029	-0.003		-0.001	
London	0.013	-0.017		-0.017		0.012	-0.018		-0.020*	
Midlands	0.060**	-0.027**		-0.025**		0.059**	-0.027**		-0.024**	
North West	0.024	-0.023**		-0.022*		0.024	-0.023**		-0.022*	
North	0.015	-0.051**		-0.048**		0.015	-0.051**		-0.047**	
Scotland	0.020	0.000		0.002		0.020	0.000		0.003	
Wales	0.073*	-0.004		-0.002		0.072*	-0.004		-0.001	
Degree		0.136**		0.126**			0.134**		0.119**	
Ahnd		0.108**		0.102**			0.107**		0.099**	
Ocse		0.090**		0.086**			0.089**		0.086**	
Nonwhite		-0.071**		-0.069**			-0.071**		-0.069**	
y1992	-0.004	-0.002		-0.002		-0.004	-0.001		-0.001	
y1993	-0.021**	-0.016**		-0.016**		-0.022**	-0.015*		-0.015**	
y1994	-0.036**	-0.028**		-0.028**		-0.037**	-0.026**		-0.027**	
y1995	-0.049**	-0.036**		-0.035**		-0.050**	-0.034**		-0.035**	
y1996	-0.068**	-0.052**		-0.051**		-0.069**	-0.049**		-0.050**	
y1997	-0.074**	-0.057**		-0.056**		-0.075**	-0.054**		-0.055**	
y1998	-0.083**	-0.064**		-0.063**		-0.084**	-0.061**		-0.061**	
y1999	0.028**	0.054**		0.054**		0.027**	0.056**		0.056**	
y2000	-0.106**	-0.080**		-0.078**		-0.108**	-0.077**		-0.077**	
y2001	-0.107**	-0.076**		-0.075**		-0.108**	-0.075**		-0.074**	
Deprivation (mundlak)									-0.029**	
cons	0.597**	0.599**		0.416**		0.569**	0.661**		0.639**	

** significant at 5%
* significant at 10%
p = pass, f = fail Hausman test

Table 3: GHQ fixed and random effects models: Women

	1	2	3	4	5	6	7	8	9	10
	FE	RE	H	REm1	H	FE	RE	H	REm1	H
Log income	0.173**	0.212**	p	0.199**	p	0.201**	0.172**	p	0.163**	p
Relative deprivation						0.030	-0.042	f	0.015	p
No. kids	0.063	0.000	p	-0.020	f	0.062	0.001	p	-0.011	f
Household size	-0.062	-0.020	p	-0.018	p	-0.063	-0.019	p	-0.015	f
Self employed	-0.042	-0.099	p	-0.017	p	-0.044	-0.102	p	-0.022	p
Unemployed	-1.644**	-1.790**	f	-1.627**	p	-1.646**	-1.790**	f	-1.636**	p
Retired	-0.355**	-0.358**	p	-0.381**	p	-0.357**	-0.357**	p	-0.400**	p
Maternity leave	-0.810**	-0.834**	p	-0.772**	p	-0.812**	-0.834**	p	-0.779**	p
Family carer	-0.771**	-0.798**	p	-0.756**	p	-0.772**	-0.798**	p	-0.771**	p
Student	0.074	-0.187	f	0.082	p	0.073	-0.189	f	0.057	p
Disabled	-2.494**	-3.576**	f	-2.499**	p	-2.494**	-3.575**	f	-2.511**	p
Other	-0.249	-0.199	p	-0.254	p	-0.251	-0.198	p	-0.267	p
Widowed	-1.618**	-1.095**	f	-1.131**	f	-1.618**	-1.093**	f	-1.118**	f
Divorced/separated	-0.849**	-1.087**	f	-0.995**	p	-0.852**	-1.083**	f	-0.983**	p
Never married	0.296**	0.253**	p	0.314**	p	0.291**	0.256**	p	0.319**	p
age3039	-0.045	-0.232**		-0.225**		-0.048	-0.235**		-0.241**	
age4049	0.071	-0.247**		-0.244**		0.066	-0.249**		-0.264**	
age5059	0.386*	0.136		0.068		0.381*	0.136		0.045	
age6069	0.919**	0.786**		0.534**		0.913**	0.790**		0.516**	
age70	0.896**	0.673**		0.281		0.891**	0.678**		0.269	
South West	0.183	0.235		0.238		0.183	0.236		0.256	
London	0.606**	-0.027		-0.003		0.607**	-0.031		-0.020	
Midlands	0.668**	0.194		0.223*		0.669**	0.196		0.247*	
North West	0.155	-0.043		0.021		0.154	-0.042		0.034	
North	0.124	-0.113		-0.082		0.122	-0.110		-0.054	
Scotland	1.725**	0.153		0.214*		1.726**	0.155		0.235*	
Wales	1.099**	-0.131		-0.048		1.100**	-0.129		-0.028	
Degree		0.901**		0.783**			0.891**		0.669**	
Ahnd		0.573**		0.457**			0.568**		0.394**	
Ocse		0.591**		0.487**			0.590**		0.460**	
Nonwhite		0.020		0.043			0.020		0.063	
y1992	-0.339**	-0.309**		-0.321**		-0.345**	-0.301**		-0.320**	
y1993	-0.490**	-0.437**		-0.445**		-0.495**	-0.428**		-0.442**	
y1994	-0.528**	-0.476**		-0.478**		-0.534**	-0.464**		-0.475**	
y1995	-0.790**	-0.717**		-0.714**		-0.798**	-0.704**		-0.710**	
y1996	-0.825**	-0.737**		-0.729**		-0.835**	-0.727**		-0.724**	
y1997	-0.793**	-0.705**		-0.689**		-0.803**	-0.693**		-0.681**	
y1998	-0.774**	-0.681**		-0.661**		-0.785**	-0.667**		-0.653**	
y1999	-0.595**	-0.480**		-0.457**		-0.605**	-0.469**		-0.448**	
y2000	-0.960**	-0.837**		-0.809**		-0.970**	-0.827**		-0.798**	
y2001	-0.920**	-0.773**		-0.737**		-0.926**	-0.768**		-0.726**	
Deprivation (mundlak)									-0.203**	
cons	23.012**	22.785**		23.074**		22.796**	23.092**		23.151**	

*** significant at 5%

* significant at 10%

p=pass, f=fail Hausman test

Table 4: Hausman-Taylor Results: Women

	SAH 1	SAH 2	GHQ 1	GHQ 2
TVexogenous				
No.kids	0.016**	0.016**	0.045	0.046
Household size	-0.002	-0.002	-0.061**	-0.059*
age3039	0.032**	0.032**	-0.089	-0.090
age4049	0.066**	0.066**	-0.015	-0.015
age5059	0.088**	0.088**	0.374*	0.374*
age6069	0.121**	0.121**	1.113**	1.115**
age70	0.088**	0.089**	1.172**	1.174**
South West	0.007	0.007	0.192	0.193
London	0.007	0.007	0.246	0.244
Midlands	-0.005	-0.005	0.287*	0.288*
NorthWest	-0.015	-0.015	-0.002	-0.001
North	-0.037**	-0.036**	-0.062	-0.061
Scotland	-0.017	-0.017	0.237	0.240
Wales	-0.013	-0.013	-0.111	-0.109
1992	-0.004	-0.003	-0.331**	-0.324**
1993	-0.021**	-0.020**	-0.476**	-0.468**
1994	-0.036**	-0.035**	-0.522**	-0.512**
1995	-0.047**	-0.046**	-0.778**	-0.767**
1996	-0.065**	-0.064**	-0.810**	-0.799**
1997	-0.072**	-0.071**	-0.784**	-0.772**
1998	-0.081**	-0.079**	-0.769**	-0.755**
1999	0.033**	0.034**	-0.580**	-0.568**
2000	-0.102**	-0.100**	-0.943**	-0.932**
2001	-0.101**	-0.100**	-0.892**	-0.885**
TVendogenous				
Log income	0.012**	0.009**	0.203**	0.168**
Relative deprivation	0.004		0.038	
Self employed	0.011	0.011	-0.037	-0.036
Unemployed	-0.031**	-0.031**	-1.649**	-1.648**
Retired	-0.041**	-0.041**	-0.316**	-0.314**
Maternity leave	-0.009	-0.009	-0.810**	-0.809**
Family carer	-0.043**	-0.043**	-0.759**	-0.759**
Student	-0.030**	-0.030**	0.102	0.101
Disabled	-0.184**	-0.184**	-2.515**	-2.515**
Other	-0.032	-0.032	-0.231	-0.230
Widowed	-0.028**	-0.028**	-1.608**	-1.607**
Divorced/separated	-0.002	-0.001	-0.911**	-0.908**
Never married	-0.009	-0.009	0.294**	0.299**
TIexogenous				
Nonwhite	-0.060**	-0.060**	0.042	0.042
TIendogenous				
Degree	0.116	0.115	-0.955	-0.963
Ahnd	0.634**	0.632**	2.418**	2.397**
Ocse	0.272**	0.272**	0.559	0.555
cons	0.375**	0.403**	22.558**	22.830**

*** significant at 5%

* significant at 10%

TVexogenous: Time varying exogenous variables

TVendogenous: Time varying endogenous variables

TIexogenous: Time invariant exogenous variables

TIendogenous: Time invariant endogenous variables

Table 5: SAH fixed and random effects models: Men

	1	2	3	4	5	6	7	8
	FE	RE m1	H	RE m2	H	FE rdep	RE m1	H
Log income	0.008**	0.012**	f	0.007**	p	0.012**	0.009	p
Relative deprivation						0.004	0.002	p
No. kids	-0.005	-0.006*	p	-0.005	p	-0.005	-0.005	p
Household size	0.003	0.005**	p	0.005**	p	0.002	0.006**	p
Self employed	0.000	0.000	p	0.000	p	0.000	-0.001	p
Unemployed	-0.032**	-0.030**	p	-0.032**	p	-0.032**	-0.032**	p
Retired	-0.069**	-0.079**	f	-0.081**	p	-0.070**	-0.082**	f
Maternity leave	-0.182*	-0.164*	p	-0.167*	p	-0.183*	-0.167*	p
Family carer	-0.057**	-0.055**	p	-0.057**	p	-0.058**	-0.058**	p
Student	-0.049**	-0.047**	p	-0.049**	p	-0.049**	-0.050**	p
Disabled	-0.218**	-0.224**	f	-0.225**	f	-0.218**	-0.225**	f
Other	-0.024	-0.017	p	-0.019	p	-0.024	-0.019	p
Widowed	-0.022	-0.003	p	-0.004	p	-0.022	-0.003	p
Divorced/separated	0.003	-0.005	p	-0.004	p	0.004	-0.004	p
Never married	-0.020*	-0.005	p	-0.005	p	-0.020*	-0.005	f
age3039	0.025**	-0.003		-0.004		0.025**	-0.004	
age4049	0.039**	-0.009		-0.011		0.039**	-0.011	
age5059	0.060**	-0.005		-0.007		0.060**	-0.007	
age6069	0.059**	-0.010		-0.011		0.059**	-0.012	
age70	0.044	-0.038**		-0.038**		0.044	-0.039**	
South West	-0.066**	-0.013		-0.012		-0.066**	-0.012	
London	-0.004	-0.027**		-0.028**		-0.003	-0.028**	
Midlands	-0.039	-0.012		-0.011		-0.039	-0.010	
North West	-0.023	-0.016		-0.015		-0.023	-0.015	
North	-0.037	-0.041**		-0.040**		-0.037	-0.039**	
Scotland	-0.048	0.007		0.008		-0.047	0.008	
Wales	0.007	-0.004		-0.003		0.007	-0.002	
Degree		0.111**		0.104**			0.101**	
Ahnd		0.076**		0.073**			0.071**	
Ocse		0.063**		0.060**			0.060**	
Nonwhite		-0.040**		-0.038**			-0.038**	
y1992	-0.016**	-0.010		-0.010		-0.017**	-0.010	
y1993	-0.029**	-0.019**		-0.019**		-0.030**	-0.019**	
y1994	-0.040**	-0.027**		-0.027**		-0.042**	-0.027**	
y1995	-0.051**	-0.037**		-0.037**		-0.052**	-0.037**	
y1996	-0.063**	-0.046**		-0.046**		-0.065**	-0.046**	
y1997	-0.072**	-0.053**		-0.052**		-0.074**	-0.052**	
y1998	-0.088**	-0.065**		-0.065**		-0.090**	-0.065**	
y1999	0.003	0.029**		0.029**		0.002	0.029**	
y2000	-0.116**	-0.085**		-0.084**		-0.117**	-0.084**	
y2001	-0.109**	-0.074**		-0.073**		-0.109**	-0.073**	
Deprivation (mundlak)							-0.018**	
cons	0.731**	0.689**		0.560**		0.698**	0.698**	

*** significant at 5%
* significant at 10%
p=pass, f=fail Hausman test

Table 6: GHQ fixed and random effects models: Men

	1	2	3	4	5	6	7	8	9	10
	FE	REm1	H	REm2	H	FE	RE	H	REm1	H
Log income	0.003	0.064*	f	-0.009	p	-0.003	0.015	p	0.009	p
Relative deprivation						-0.006	-0.072	f	-0.055	p
No. kids	0.011	-0.012	p	-0.003	p	0.012	-0.029	p	-0.011	p
Household size	-0.018	0.005	p	0.008	p	-0.018	0.023	p	0.008	p
Self employed	-0.096	-0.074	f	-0.083	p	-0.092	-0.106	p	-0.071	p
Unemployed	-1.973**	-1.942**	f	-1.962**	p	-1.973**	-2.160**	f	-1.941**	f
Retired	-0.260**	-0.300**	p	-0.332**	f	-0.259**	-0.537**	f	-0.295**	p
Maternity leave	-3.100**	-2.829**	p	-2.870**	p	-3.097**	-2.491**	p	-2.819**	p
Family carer	-0.803**	-0.768**	p	-0.810**	p	-0.801**	-0.971**	p	-0.766**	p
Student	-0.405**	-0.355**	p	-0.394**	p	-0.405**	-0.435**	p	-0.358**	p
Disabled	-2.732**	-2.760**	p	-2.783**	f	-2.732**	-3.827**	f	-2.758**	p
Other	-0.991**	-0.963**	p	-0.983**	p	-0.988**	-1.095**	p	-0.960**	p
Widowed	-0.952**	-0.758**	p	-0.761**	p	-0.952**	-0.738**	p	-0.756**	p
Divorced/separated	-1.196**	-0.966**	f	-0.947**	f	-1.197**	-1.055**	p	-0.968**	f
Never married	-0.207*	0.001	f	0.006	f	-0.206*	0.000	f	0.004	f
age3039	-0.258**	-0.590**		-0.605**		-0.258**	-0.643**		-0.593**	
age4049	-0.408**	-0.876**		-0.898**		-0.405**	-0.983**		-0.876**	
age5059	-0.055	-0.483**		-0.506**		-0.052	-0.662**		-0.482**	
age6069	0.617**	0.193		0.171		0.620**	0.080		0.197	
age70	0.524	-0.236		-0.248		0.527	-0.294*		-0.232	
South West	0.408	0.294*		0.310**		0.410	0.311**		0.297**	
London	-0.002	-0.029		-0.041		0.004	-0.062		-0.031	
Midlands	0.721**	0.236*		0.263**		0.722**	0.226*		0.239*	
North West	0.621*	0.269*		0.280*		0.623*	0.213		0.271*	
North	0.795**	0.270**		0.290**		0.797**	0.227*		0.274**	
Scotland	0.200	0.383**		0.400**		0.200	0.309**		0.385**	
Wales	0.760*	0.101		0.117		0.761*	-0.015		0.104	
Degree		-0.215		-0.337**			0.024		-0.230*	
Ahnd		-0.040		-0.105			0.170		-0.045	
Ocse		0.078		0.035			0.249**		0.076	
Nonwhite		-0.300		-0.265			-0.295		-0.301	
y1992	-0.251**	-0.174**		-0.172**		-0.249**	-0.125		-0.161*	
y1993	-0.152*	-0.056		-0.054		-0.149*	0.003		-0.042	
y1994	-0.261**	-0.138		-0.135		-0.258**	-0.072		-0.122	
y1995	-0.342**	-0.211**		-0.206**		-0.339**	-0.135		-0.192**	
y1996	-0.418**	-0.265**		-0.257**		-0.415**	-0.188**		-0.247**	
y1997	-0.395**	-0.232**		-0.221**		-0.392**	-0.154*		-0.212**	
y1998	-0.410**	-0.216**		-0.204**		-0.406**	-0.133		-0.192**	
y1999	-0.418**	-0.213**		-0.202**		-0.415**	-0.129		-0.193**	
y2000	-0.718**	-0.464**		-0.450**		-0.715**	-0.379**		-0.446**	
y2001	-0.739**	-0.466**		-0.453**		-0.737**	-0.385**		-0.456**	
cons	25.990**	26.006**		23.694**		26.033**	26.052**		26.433**	

*** significant at 5%

* significant at 10%

p=pass, f=fail Hausman test

Table 7: Hausman-Taylor Results: Men

	SAH 1	SAH 2	GHQ 1	GHQ 2
TVexogenous				
No.kids	-0.006	-0.006	-0.007	-0.007
Household size	0.003	0.004	-0.005	-0.005
age3039	0.020**	0.020**	-0.435**	-0.435**
age4049	0.033**	0.033**	-0.632**	-0.632**
age5059	0.054**	0.054**	-0.198	-0.198
age6069	0.064**	0.064**	0.597**	0.597**
age70	0.060**	0.061**	0.448	0.448
South West	-0.035**	-0.035**	0.353*	0.353*
London	0.001	0.000	-0.022	-0.023
Midlands	-0.016	-0.016	0.370**	0.370**
NorthWest	-0.025	-0.025	0.304	0.304
North	-0.047**	-0.046**	0.371*	0.371**
Scotland	-0.004	-0.004	0.323*	0.323*
Wales	-0.023	-0.023	0.080	0.080
1992	-0.016**	-0.015	-0.217**	-0.217**
1993	-0.029**	-0.028**	-0.117	-0.117
1994	-0.040**	-0.039**	-0.216**	-0.216**
1995	-0.051**	-0.050**	-0.296**	-0.296**
1996	-0.063**	-0.061**	-0.361**	-0.361**
1997	-0.073**	-0.071**	-0.338**	-0.338**
1998	-0.089**	-0.086**	-0.338**	-0.338**
1999	0.004	0.006	-0.345**	-0.345**
2000	-0.113**	-0.111**	-0.613**	-0.613**
2001	-0.105**	-0.104**	-0.629**	-0.629**
TVendogenous				
Log income	0.013**	0.008**	-0.004	-0.004
Relative deprivation	0.006		0.0004	
Self employed	-0.001	-0.001	-0.102	-0.102
Unemployed	-0.033**	-0.033**	-1.968**	-1.968**
Retired	-0.065**	-0.065**	-0.240**	-0.240**
Maternity leave	-0.185**	-0.184**	-3.110**	-3.110**
Family carer	-0.060**	-0.060**	-0.810**	-0.810**
Student	-0.048**	-0.048**	-0.330**	-0.330**
Disabled	-0.221**	-0.221**	-2.740**	-2.740**
Other	-0.024	-0.024	-0.974**	-0.974**
Widowed	-0.018	-0.018	-0.993**	-0.993**
Divorced/separated	0.000	-0.001	-1.237**	-1.237**
Never married	-0.01*	-0.018*	-0.111	-0.111
TIexogenous				
Nonwhite	0.000	0.000	-0.370	-0.370
TIendogenous				
Degree	0.112	0.112	1.784	1.784
Ahnd	0.539**	0.537**	2.361**	2.361**
Ocse	0.435**	0.434**	1.223	1.223
cons	0.398**	0.438**	25.014**	25.017**

*** significant at 5%

* significant at 10%

TVexogenous: Time varying exogenous variables

TVendogenous: Time varying endogenous variables

TIexogenous: Time invariant exogenous variables

TIendogenous: Time invariant endogenous variables

Table 8: Semiparametric Results

	Women: GHQ	Women: SAH	Men:GHQ	Men: SAH
Relative deprivation	-0.162	-0.002	-0.408**	0.012
No. kids	-0.005**	0.017	0.001	-0.005
Household size	-0.020	-0.001	0.008	0.006**
Self employed	-0.018	0.011	-0.069	0.000
Unemployed	-1.628**	-0.030**	-1.935**	-0.030**
Retired	-0.392**	-0.042**	-0.310**	-0.080**
Maternity leave	-0.757	-0.004**	-2.718**	-0.166*
Family carer	-0.762**	-0.043**	-0.771**	-0.055**
Student	0.059**	-0.030	-0.365**	-0.049**
Disabled	-2.511**	-0.187**	-2.772**	-0.223**
Other	-0.255	-0.034	-0.949**	-0.018
Widowed	-1.119**	-0.025**	-0.764**	-0.004
Divorced/separated	-0.967	-0.006**	-0.948**	-0.004
Never married	0.321**	0.023**	0.006	-0.005
age3039	-0.241	0.004**	-0.606**	-0.004
age4049	-0.266**	0.016**	-0.896**	-0.011
age5059	0.045**	0.023	-0.503**	-0.008
age6069	0.514**	0.038**	0.175	-0.012
age70	0.265	-0.011	-0.250	-0.038*
South West	0.255	-0.001	0.323**	-0.012
London	-0.018*	-0.020	-0.047	-0.028**
Midlands	0.249**	-0.024*	0.275**	-0.010
North West	0.040*	-0.021	0.282*	-0.015
North	-0.046**	-0.047	0.307**	-0.039**
Scotland	0.241	0.003*	0.409**	0.008
Wales	-0.021	-0.001	0.130	-0.002
Degree	0.679**	0.119**	-0.374**	0.101**
Ahnd	0.397**	0.100**	-0.111	0.071**
Ocse	0.456**	0.086**	0.033	0.059**
Nonwhite	0.075**	-0.069	-0.266	-0.038**
y1992	-0.277	-0.001**	-0.065	-0.013
y1993	-0.395*	-0.015**	0.060	-0.022**
y1994	-0.421**	-0.027**	0.001	-0.030**
y1995	-0.646**	-0.034**	-0.045	-0.041**
y1996	-0.662**	-0.050**	-0.102	-0.050**
y1997	-0.611**	-0.054**	-0.042	-0.057**
y1998	-0.576**	-0.061**	-0.003	-0.070**
y1999	-0.385**	0.056**	-0.031	0.025**
y2000	-0.741**	-0.077**	-0.293**	-0.089**
y2001	-0.702**	-0.074**	-0.374**	-0.076**
constant	-0.020	-0.002	-0.047	-0.006**

*** significant at 5%

* significant at 10%

References

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

Blanchflower, D. and Oswald, A. (2004). Well-being overtime in Britain and the USA. *Journal of Public Economics*, 88:1359–1386.

Bowling, A. (1991). *Measuring Health. A review of quality of life measurement scales.* OUP, Milton Keynes, first edition.

Chakravarty, S. (1990). *Social Ethical Index Numbers.* Springer-Verlag, Berlin.

Clark, A. (2003). Inequality-aversion and income mobility: A direct test. CNRS and DELTA-Fédération Jourdan, mimeo.

Clark, A. and Oswald, A. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61:359–381.

Contoyannis, P., Jones, A. M., and Rice, N. (2003). The dynamics of health in the British Household Panel Survey. forthcoming in the *Journal of Applied Econometrics.*

Deaton, A. (2003). Health, inequality and economic development. *Journal of Economic Literature*, 41:113–158.

Desai, M. and Shah, A. (1988). An econometric approach to the measurement of poverty. *Oxford Economic Papers*, 40:505–522.

Duesenberry, J. S. (1949). *Income, Saving and the Theory of Consumer Behaviour.* Harvard University Press, Cambridge MA., first edition.

Ecob, R. and Davey Smith, G. (1999). Income and health: what is the nature of the relationship? *Social Science and Medicine*, 48:693–705.

Estaban, J. and Ray, D. (1994). On the measurement of polarization. *Econometrica*, 62:819–851.

Ettner, S. (1996). New evidence on the relationship between income and health. *Journal of Health Economics*, 15:67–85.

Frank, R. (1985a). The demand for unobservable and other nonpositional goods. *American Economic Review*, 75:101–116.

Frank, R. (1997). The frame of reference as a public good. *Economic Journal*, 107:1832–1847.

Frank, R. H. (1985b). *Choosing the Right Pond.* Oxford Univeristy Press, New York.

Goldberg, D. and Williams, P. (1988). *A User's Guide to the General Health Questionnaire.* Nfer-Nelson, Windsor.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–55.

Härdle, W. (1990). *Applied Nonparametric Regression.* Cambridge University Press, Cambridge.

Hausman, J. and Taylor, W. (1981). Panel data and unobservable individual effects. *Econometrica*, 49:1377–1398.

Hernandez-Quevedo, C., Jones, A. M., and Rice, N. (2004). Reporting bias and heterogeneity in assessed health. evidence from the british household panel survey. University of York, mimeo.

Hey, J. and Lambert, P. (1980). Relative deprivation and the Gini coefficient: comment. *Quarterly Journal of Economics*, 95:567–573.

Likert, R. (1952). A technique for the development of attitude scales. *Educational and Psychological Measurement*, 12:313–315.

Lindley, J. and Lorgelly, P. (2003). The relative income hypothesis: does it exist over time? Evidence from the BHPS. University of East Anglia, mimeo.

Lynch, J., Davey Smith, G., Harper, S., Hillemeier, M., Ross, N., Kaplan, G., and Wolfson, M. (2004). Is income inequality a determinant of population health? Part 1. A systematic review. *Millbank Quarterly*, 82:5–99.

Marmot, M., Rose, G., Shipley, M., and Hamilton, P. (1978). Employment grade and coronary heart disease in British civil servants. *Journal of Epidemiology and Community Health*, 32:244–249.

Mundlak, Y. (1978). On the pooling of time series and cross-section data. *Econometrica*, 46:69–85.

Robinson, P. M. (1988). Root-N consistent semiparametric regression. *Econometrica*, 56:931–954.

Runciman, W. (1966). *Relative Deprivation and Social Justice*. Routledge and Kegan Paul, London, first edition.

Sapolski, R. (1993). Endocrinology alfresco: Psychoendocrine studies of wild baboons. *Recent Progress in Hormone Research*, 48:437–468.

Soobader, M. and LeClere, F. (1999). Aggregation and the measurement of income inequality: effects on morbidity. *Social Science and Medicine*, 48:733–744.

Townsend, P. (1979). *Poverty in the United Kingdom*. Allen Lane and Penguin, London.

van Doorslaer, E. and Gerdtham, U.-G. (2003). Does inequality in self-assessed health predict inequality in survival by income? - Evidence from Swedish data. *Social Science and Medicine*, 57:1621–1629.

van Doorslaer, E., Wagstaff, A., and et al. (1997). Income-related inequalities in health: some international comparisons. *Journal of Health Economics*, 16:93–112.

Wildman, J. (2003). Modelling health, income and income inequality: the impact of income inequality on health and health inequality. *Journal of Health Economics*, 22:521–538.

Wildman, J. and Jones, A. M. (2001). Inequalities in health. In Askildsen, J. and Haug, K., editors, *Helse, økonomi og politikk: utfordringer for det norske helsevenet*, pages 37–62. Cappelen, Oslo.

Wilkinson, R. (1996). *Unhealthy Societies: The Afflictions of Inequality.* Routledge, London.

Wooldridge, J. (2002). *Econometric Analysis of cross section and panel data.* MIT press, Cambridge, Massachusetts.

Figure 1: Density of log income and Smoothed function of GHQ and income: women

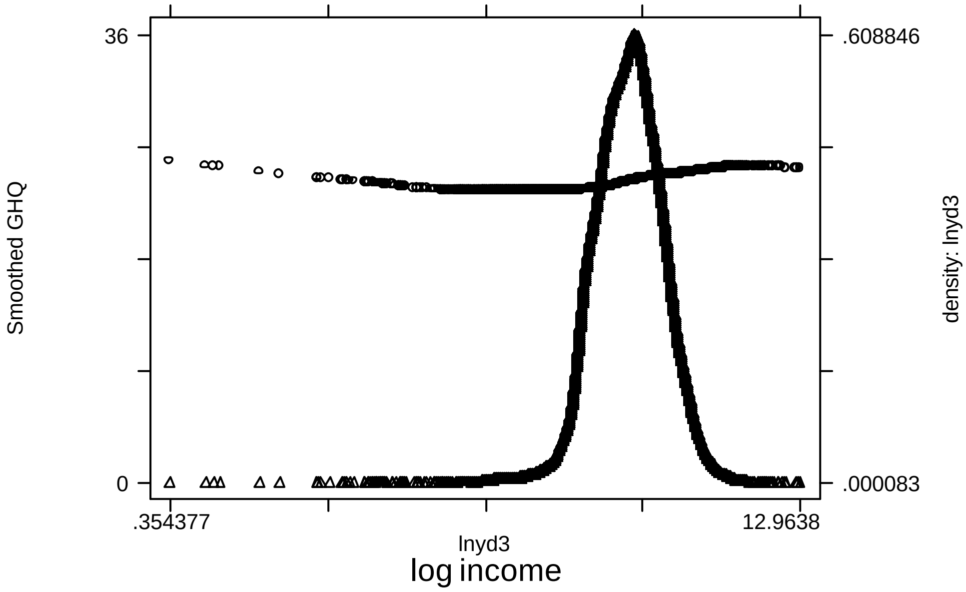

Figure 2: Density of log income and Smoothed function of SAH and income: women

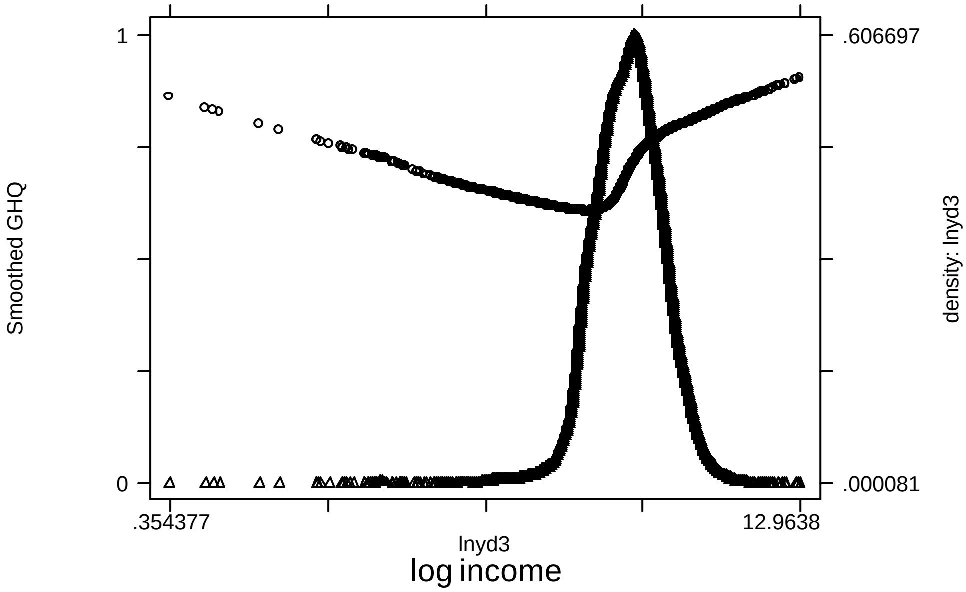

Figure 3: Density of log income and Smoothed function of GHQ and income: men

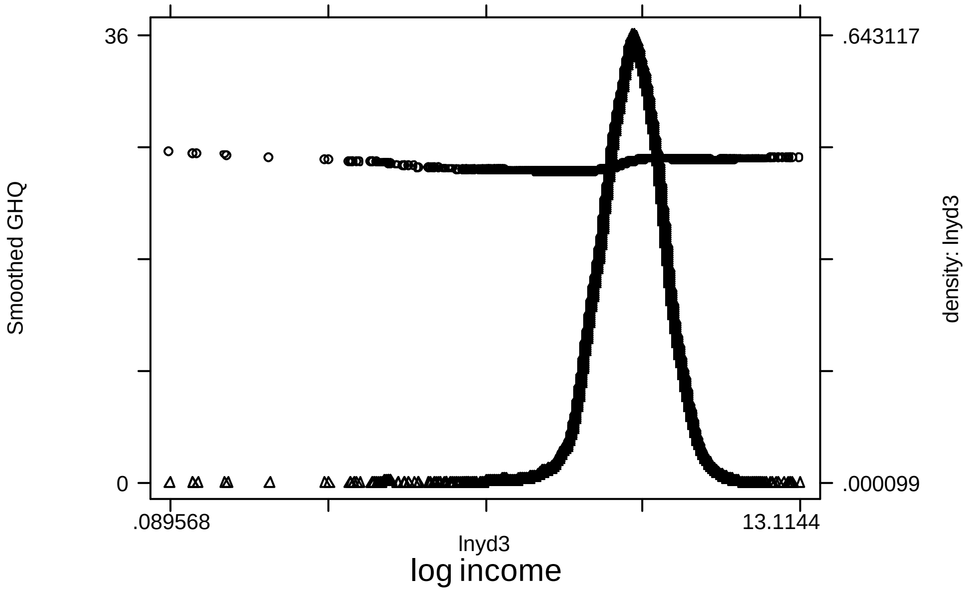

Figure 4: Density of log income and Smoothed function of SAH and income: men