Health and Human Capital

Mary O'Mahony* and Lea Samek*

Abstract

This study examines the impact of morbidity on human capital stocks (HCS) with an application to the UK from 1996 to 2018. It incorporates health status into the standard Jorgenson-Fraumeni lifetime income measure of HCS through its effect on absenteeism and presenteeism (lost productivity) by taking account of inactivity due to illness and modelling the impact of health on earnings and retirement behaviour. It employs the approach of estimating individual health indices, which reduce concerns for reporting and errors-in-variable bias, and takes account of individuals' and spouses' health as well as caring responsibilities due to adverse health of third parties. The results show that overall poor health leads to a reduction in HCS by about 12 per cent in 2018, but shows a slight tendency to decrease over time. This is mostly driven by inactivity due to long-term illness, and retirements for those aged over 50. There are significant impacts of poor health on earnings, especially for males, but the results show only a small impact on HCS from earnings as most people in poor health are not economically active. The results vary by qualification level, gender and age, with productive HCS reduced by about 45 per cent for individuals aged 50 years or older with low qualifications.

* King's College London and Economic Statistics Centre of Excellence (ESCoE)

1. Introduction

Although it is well known that physical health status plays a significant role in generating effective labour supply, existing measures of *human capital stocks* (HCS) are inadequate in dealing with health status. Poor health can have a direct impact on the quantity of labour supplied, through short-term absence from work, inactivity or permanent retirement, and, in the extreme, through mortality. It can also reduce on-the-job productivity and work quality, referred to as *presenteeism* in the occupational health literature (Burton *et al.* 1999; Turpin *et al.* 2004)[1], which imposes high costs often exceeding those incurred through absenteeism (Garrow, 2016) or through direct medical care (Loeppke *et al.* 2003; Goetzel *et al.* 2004). These health effects may even spill over to workers in good health because their work routine is likely to be impacted.

A commonly used method by academics and statistical offices to measure HCS is that developed by Jorgenson and Fraumeni (JF) (1989, 1992). In the JF framework lifetime income is used to measure HCS, which sums the discounted values of all future income streams that the population is expected to earn throughout their lifetime. Applications of this method take account of health explicitly only in its most extreme manifestation, mortality. Many applications measure the '*productive'* HCS of the workforce by allowing for employment rates or hours worked rather than focusing on the total population. To the extent that early retirement or days lost due to absence are due to ill-health, these measures implicitly take account of health status. Presenteeism is not generally captured, except to the extent that it affects earnings. Explicitly modelling and estimating the relationship between health and HCS is required to estimate the costs of ill-health to the economy through its impact on effective labour supply. This paper is a first attempt to bring together the literatures on health and human capital in a unified framework which can then be used, e.g. to evaluate cost effectiveness of interventions to improve health conditions. It first sets out the model in section 2 and then implements this using data for the UK.

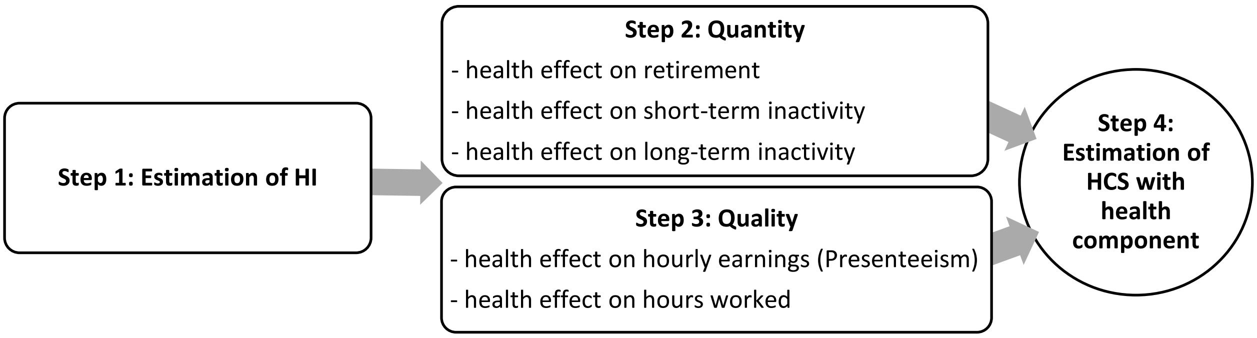
In this paper we make a distinction between HCS for the active population (excluding those

[1] A second literature strand has emerged that focuses on individuals attending work while ill which is discussed in more detail by Johns (2010).

who are very young or very old) and for the employed population. Implementing the framework for the UK requires a number of steps, as illustrated in Figure 1, and detailed in sections 3 and 4. This involves first defining what we mean by health status, which in itself is a difficult concept to implement. We then model explicitly how health impacts on decisions to retire from the workforce, reducing the active HCS, and on earnings, reducing the employed HCS. These health impacts are estimated using longitudinal data from the Understanding Society Survey or the United Kingdom Household Longitudinal Study (UKHLS). We then incorporate these impacts into measures of HCS using data from the Labour Force Survey (LFS), the primary source used in the current ONS estimates.

Standard in the literature, we use a self-constructed health index (HI) as follows. A categorical self-assessed health (SAH) variable is regressed on measures of diagnosed health conditions, a number of functional limitations, a mental health measure and a broad range of socio-economic characteristics using generalised ordered probit regression. Therefore, we exploit the advantages of subjective and objective health indicators and account for prevalence as well as severity of specific health conditions. Health adjusted employment rates for older age groups are then constructed by estimating the effect of individual and spouse's health as well as caring responsibilities on the probability to retire. Absenteeism is implicitly accounted for in the annual earnings data through lower annual working hours and both absenteeism and presenteeism affect wages. We estimate the impact of health status on hourly wages for those who are employed, and translate predicted wages into annual earnings using hours worked. The specifications and the econometric and data issues involved in these estimations are outlined in section 3. Section 4 then brings together the analysis from the previous sections to estimate the impact of health on HCS and section 5 concludes.

Figure 1: Overview of Research Approach



The results in section 4.2 show that overall poor health leads to a reduction in HCS by about 13% averaged across the time period 1996-2018, with a small decline across time, from 14% in the late 1990s to 12% today. This is mostly driven by inactivity due to long-term illness (9 percentage points). Short-term illnesses reduce HCS by 2 percentage points. There is also an impact from presenteeism and absenteeism (including early retirement) but that is also relatively small (1 percentage point each). These health impacts are marginally higher for females than for males. HCS is reduced by about 45 per cent for individuals aged 50 or older in poor health with low qualifications.

2. Including Health in the Jorgenson Fraumeni Framework

The modern approach in assessing the importance of human capital in driving economic growth and the resulting policy implications stem from the work by Shultz (1961), Becker (1964) and Mincer (1974) but the importance of health in this context was only recognised since the mid-1990s (e.g. Barro, 1996; Lopez-Casanovas *et al.* 2005, among others). Moreover, human capital is complex and encompasses "knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being" (Dang *et al.* 2001). Therefore, it is difficult to capture all aspects in one model and measurement tended to focus on single aspects, especially education.

There are essentially three methods employed to measure HCS, cost-based, educational attainment-based or income-based (see Jones and Fender, 2011 and Samek, 2017 for reviews). The cost-based measure stems from the seminal work of Kendrick (1976) and relies on inputs, such as education, training, health, mobility and safety. It uses a depreciated value of these monetary investments in human capital to estimate its stock. The chief problem with this approach is separating consumption from investment, e.g. how to classify child rearing expenditures. Incorporating health into this framework is feasible but complex and would not be easily broken down by level of education. The educational attainment-based estimates of human capital use indicators closely related to education. Examples include average years of schooling, adult literacy, school enrolment, or test scores as quality measures (see Hanusheck and Kimko, 2000). Many indicators, especially test scores, are difficult to compare across time and countries. In theory these could be combined with health status indicators, but it is not clear what weights could be used when measuring the stock of human capital.

A popular method used by many countries is to consider outcomes from human capital accumulation using information on lifetime incomes. The seminal contribution to this literature were the papers by Jorgenson and Fraumeni (1989, 1992). The JF model measures HCS using lifetime earnings in present discounted value that all individuals are expected to earn. This implies the assumption that labour is paid according to its marginal product. The JF approach is consistent with the treatment of assets in the national accounts (Fraumeni *et al.* 2017) and is the approach recommended by the Atkinson Report (2005). We chose to employ the income-based approach as it allows the incorporation of health as an additional capital driver in a labour market context. By adjusting earnings and labour force participation accordingly, our method values the productive capacity of the workforce after taking account of health status. There are a large number of international efforts to measuring human capital accounts using the JF framework[2], but none take account of morbidity aspects. Although it can be argued that health is implicitly accounted for in the JF lifetime income framework through the employment and survival rate of the employed population, these estimations fail to make the additional impact of health transparent by singling out the distinctive effect of morbidity. Moreover, this study focuses on morbidity rather than mortality by studying labour potentially available for production and how its availability changes with health.

[2] For example: see Ahlroth *et al.* (1997) for an early application to Sweden, and subsequently Wei (2004) for Australia; Gu and Wong (2010) for Canada; Jones and Chiripanhura (2010) for the UK; and Christian (2010) for the US, to name but a few.

In its original formulation the JF framework calculates lifetime income by sex (*s*), age(*a*), and education level (*e)* and then sums across these dimensions for the total population. In this paper we confine attention to the *'potential'* HCS by focusing on the working age population, aged 16 to 69. This group can be divided into five categories, those in employment, unemployment, retirement, absent due to sickness and economic inactivity other than sickness. For individuals not in employment, we impute potential earnings based on employment earnings for similar demographic groups.[3]

Lifetime labour income, $LLI_{s,a,e}$, is calculated by using backwards recursion. This implies that market income is zero beyond some age, which in this paper is taken to be 69, and is based on the assumption that people do not receive any earnings once they withdraw from the labour market.

Therefore, lifetime earnings of those aged 69 is given by:

$$LLI_{s,a=69,e} = income_{s,a=69,e} \tag{1}$$

The JF methodology assumes that an individual with a given gender, age and education will, in year t+1, have the same labour income and other characteristics (employment and survival rate) as someone in year t, who is one year older and has otherwise the same characteristics (gender and education). Therefore, if someone is aged 68, this person´s LLI equals current income plus discounted future income of someone aged 69 with the same sex and education, conditional on survival, sr. Similar calculations apply to all persons aged above the maximum school enrolment age, which we assume equals 40.

These are given by:

$$LLI_{s,a,e} = income_{s,a,e} + sr_{s,a+1}\frac{1+g}{1+\delta}LLI_{s,a+1,e} \tag{2}$$

$$|\ 40 \leq a \leq 69$$

where g equals the labour productivity growth rate and δ represents the discount rate.

[3] In the original JF framework a value based on the average market wage was also imputed to nonworking time after adjusting for maintenance (time spent sleeping, eating etc.). This controversial assumption was not widely adopted in subsequent estimates and is not included here, but see Fraumeni *et al.* (2017) for a recent effort to integrate this broader measure into the US national accounts.

For those aged between 16 and 39, LLI needs to take account of education enrolment (school, further education and higher education[4]), $ENRR$[5]. Therefore, equation (2) is altered to include the probability of people improving their educational attainment, which is multiplied by the income they are likely to earn given their higher qualification. At the start of each year, everyone has the choice to either work next year maintaining the same qualification level, $(1 - ENRR_{s,a,e})LLI_{s,a+1,e}$ or improve it and, hence, receive a different income, $ENRR_{s,a,e}LLI_{s,a+1,e+1}$:

$$LLI_{s,a,e} = income_{s,a,e} + sr_{s,a+1}\frac{1+g}{1+\delta}\left[ENRR_{s,a,e}LLI_{s,a+1,e+1} + (1 - ENRR_{s,a,e})LLI_{s,a+1,e}\right] \\ \mid 16 \leq a \leq 39 \tag{3}$$

Total HCS is calculated by aggregating individual LLI across the population, POP:

$$potential\ HCS = \sum_s \sum_a \sum_e LLI_{s,a,e}POP_{s,a,e} \tag{4}$$

The calculations for the productive HCS are similar to equations (2) – (4) except we now multiply incomes by the employment rate, $EMPR$:

$$LLI_{s,a,e} = EMPR_{s,a,e}income_{s,a,e} + sr_{s,a+1}\frac{1+g}{1+\delta}LLI_{s,a+1,e} \\ \mid 40 \leq a \leq 68 \tag{5}$$

$$LLI_{s,a,e} = \\ EMPR_{s,a,e}income_{s,a,e} + sr_{s,a+1}\frac{1+g}{1+\delta}\left[ENRR_{s,a,e}LLI_{s,a+1,e+1} + (1 - ENRR_{s,a,e})LLI_{s,a+1,e}\right] \\ \mid 16 \leq a \leq 39 \tag{6}$$

In this case the total productive HCS is given by:

$$productive\ HCS = \sum_s \sum_a \sum_e LLI_{s,a,e}EMPR_{s,a,e}POP_{s,a,e} \tag{7}$$

Calculating the productive HCS is the approach adopted in many countries, e.g. for Canada and the UK, as it measures the human capital being used in the productive process. However,

[4] These estimates are constructed as enrolment probabilities between calendar years. Since data is currently only available up to 2017/18, we assume the same enrolment rates for 2017 and 2018.
[5] Although the cut-off point is arbitrary here, actual enrolment rates do not show much education activity beyond this age.

identifying sources of the difference between potential HCS and productive HCS can be useful in understanding labour market effects, and in the case of this paper, health effects.

Incorporating health involves two components, which are estimated simultaneously. The first is that poor health reduces the potential HCS by removing some persons from the labour market. First are those who retire permanently from the labour force. In this paper we assume that retirement only affects those aged 50 and over. In addition we take account of those who explicitly state that the reason for their inactivity is ill health, either short-term or long-term. The unemployment rates are assumed to not depend on health, which is unlikely to be true but is difficult to take into account given our data. We therefore calculate the difference between potential and productive HCS and calculate how much this difference is due to poor health by modelling the impact of health on retirement, as set out in section 3.3.

Without accounting for health, the employment rate is given by:

$$EMPR_{s,a,e} = (1 - RETR_{s,a,e} - UNEMPR_{s,a,e} - SICK_{s,a,e} - INAR_{s,a,e}) \quad (8)$$

Adjusting for health, the employment rate is further divided by health status, h:

$$EMPR_{s,a,e,h} = (1 - RETR_{s,a,e,h} - UNEMPR_{s,a,e} - SICK_{s,a,e} - INAR_{s,a,e}) \quad (9)$$

where $RETR$ is the retirement and $UNEMPR$ is the unemployment rate. $SICK$ captures the proportion of people who are inactive due to sickness while $INAR$ captures everyone in economic inactivity for other reasons.

We calculate the differences between potential and productive HCS due to greater retirement or inactivity for those in poor health relative to those in good health. It measures how much the productive HCS is reduced by persons leaving the workforce prematurely due to poor health. This we call the '*quantity health effect*'. Incorporating health into the potential and productive HCS framework is straightforward, by further dividing by health status. For the productive HCS, and those aged 35 to 69, equation (5) above is replaced by:

$$LLI_{s,a,e,h} = EMPR_{s,a,e,h} income_{s,a,e,h} + sr_{s,a+1} \frac{1+g}{1+\delta} LLI_{s,a+1,e,h} \quad (10)$$

$$| \; 40 \leq a \leq 69$$

For those aged between 16 and 39, equation (6) is replaced by:

$$LLI_{s,a,e,h} = EMPR_{s,a,e,h} income_{s,a,e,h} +$$
$$sr_{s,a+1} \frac{1+g}{1+\delta} \left[ENRR_{s,a,e} LLI_{s,a+1,e+1,h} + \left(1 - ENRR_{s,a,e}\right) LLI_{s,a+1,e,h}\right]$$
$$| \ 16 \leq a \leq 39 \qquad (11)$$

Similarly equations (2) and (3) distinguish health status for the potential HCS. Total health adjusted productive HCS is calculated by aggregating individual LLI in each age and qualification category across the employed, and then summing by gender and health to make the health effect transparent.

We then calculate how much this productive HCS would rise if all those in poor health were in fact in good health. We call this the *'quality health effect'*. Finally, in order to implement the above model we need to divide the population/employment according to health status and estimate the extent to which employment and wages vary by health. In this paper we assume that poor health only affects those aged 26 and over. The next section addresses these issues, before we turn to populating the model.

3. Impacts of Health on Labour Supply and Wages

The following section describes the data sources and methods used to construct the HI and explains how it is used in the estimation of health adjusted labour supply and wages. We distinguish between retirement and absenteeism when examining labour supply.

3.1 Data

The analyses reported in this section uses data from all eight waves of the UKHLS, which is a longitudinal survey based on a representative probability sample of approximately 40,000 households from England, Scotland, Wales and Northern Ireland. It includes questions on health, work, education, income, family and social life, and offers the opportunity to follow approximately 6,900 original sample members from the British Household Panel Survey, who became part of the UKHLS. To be eligible for full interviews, household members had to be aged 16 or older and interviews were conducted yearly face-to-face in their homes. Data collection started in 2009 and took place over a 24-month period for each wave. Consequently, each wave does not necessarily correspond to the same calendar year across all participants and some waves overlap in particular calendar years.

3.2 Constructing a Health Index

When measuring individual health status, early research often relied on objective health measures (e.g. Stern, 1989; Anderson and Burkhauser, 1985), but later work questioned the validity of just relying on these measures as they are unlikely to be perfectly correlated with the health aspect that affects work capabilities (e.g. Bound *et al.* 1999; Hernández-Quevedo *et al.* 2005). Also, they are likely to be collinear and they ignore the severity of the disease by describing someone's health only partially. At the same time, using them separately in the model makes it difficult to quantify the marginal effect of health changes. Therefore, research turned to quantifying health effects using self-assessed health (SAH) (e.g. Anderson and Burkhauser, 1985; Stern, 1989; Bound *et al.* 1999; van Doorslaer and Jones, 2003; Disney *et al.* 2006; Lindeboom and Kerkhofs, 2009; Jones *et al.* 2010). Many general population surveys, such as the UKHLS, include this variable by asking respondents to rate their own health on a categorical scale. Although it does not capture an individual's underlying health stock perfectly, its predictive power of subsequent health status has been widely researched (e.g. Idler and Kasl, 1995; Idler and Benyamini, 1997; van Doorslaer and Gerdtham, 2003) and it inherits a sense of ordering. Although its subjective nature captures the severity of a potential health condition, at the same time this compromises its validity since individuals with similar health may use different thresholds when assessing their own health. This '*threshold heterogeneity'* can be influenced by age, education, employment or personal experiences among many other influences and is likely to bias the estimated health effect on the variable of interest (Bound, 1991). Generally, health can be written as a function of objective health indicators and other control variables, x_{it}:

$$h_{it} = \beta' x_{it} + \varepsilon_{it}, \tag{13}$$
$$i = 1, 2, \dots, n; t = 1, 2, \dots, T$$

where h_{it} represents health of an individual *i* at time *t* and ε_{it} is the time varying error term, which is uncorrelated with x_{it}.

True health, h_{it}, however, cannot be observed directly so the categorical variable SAH is frequently used as a proxy. When studying the aspect of health that impacts on the quantity as well as the quality of labour, the potential measurement error in SAH has to be addressed. Following Disney *et al.* (2006), who use the approach set out in Bound (1991) and implemented by Bound *et al.* (1999), we construct a HI for men and women separately using eight waves from the UKHLS. SAH is estimated as a function of diagnosed health conditions[6], functional limitations[7] and a number of socio-economic indicators including employment status and household income. This procedure is analogous to using objective health measures as an instrument for the endogenous and potentially error-ridden SAH variable. Previous work relied on an ordered probit regression method which assumes constant thresholds between different SAH categories and uniform effects on every individual´s underlying health (*'parallel line assumption'*). However, either the relative position of one threshold can shift (e.g. healthy respondents report *poor* health to justify earlier retirement) or all thresholds can move in a parallel manner (e.g. a sub-group of individuals is more modest about their health status driven by their gender, age, culture, etc.), so that inferred health predictions are unlikely to reflect true health correctly. Since tests reported in Samek (2017) show that threshold shifts are evident in our data, violating the parallel line assumption[8], we combine elements from more recent work by Jones *et al.* (2010) and Pfarr *et al.* (2012a, 2012b) and rely on generalised ordered probit (GOP). We extend Jones *et al.*'s (2010) model by using functional limitations to further control for the severity rather than only the incidence of conditions.

GOP nests the standard ordered probit regression with the restriction that all threshold parameters are the same. Therefore, we refer to the former as the 'unconstrained' and to the latter as the 'constrained' regression. Given that the unconstrained model consists of a

[6] These include asthma, arthritis, congestive heart failure, coronary heart disease, angina, heart attack, stroke, emphysema, hyperthyroidism, hypothyroidism, chronic bronchitis, liver condition, cancer, diabetes, epilepsy, high blood pressure, clinical depression or any other diagnosed condition.

[7] These include difficulties with mobility, lifting/carrying/moving objects, manual dexterity, continence, hearing, sight, communication or speech, memory or ability to concentrate/learn/understand, recognising when you are in physical danger, physical co-ordination, with own personal care or some other health problem.

[8] By relaxing the parallel line assumption for objective variables exogenous to actual health, any significant variation across thresholds can be exclusively assigned to respondents' different reporting behaviour and justifies a more flexible specification than allowed for by a standard ordered probit regression. Samek (2017) finds gender-specific threshold shifts driven by personality traits and interview conditions, which are exogenous to health, but also by health conditions, functional limitations and some socio-economic characteristics. Although the justification hypothesis with respect to retirement, i.e. were health is misreported to justify non-work, is not evident, it is observed among all other economically inactive respondents, which is consistent with existing findings.

number of binary ordered probit regressions, we can test individual coefficients, first by relaxing the parallel line assumption for all covariates and then by imposing constraints in a stepwise procedure based on their significance. The least significant variable is assumed to have the same effect across all equations and, thus, is constrained to meet the parallel line assumption. The model is estimated again with the constrained variable and the process is repeated until all remaining variables are significant. The resulting specification is shown in equation (14)[9]. Results are presented in Appendix A1 and A2 for men and women, respectively.

$$\Pr\left(h_{it} \leq \mathrm{j}\middle|x_{it}^{constrained}, x_{it}^{unconstrained}\right) = \Phi\left(\tilde{\mu}_j - \beta' x_{it}^{constrained} - \delta_j x_{it}^{unconstrained}\right) \quad (14)$$

$$j = poor, fair, good, very\ good, excellent\ SAH$$

For the construction of the HI as a measure of men´s and women´s individual health, a linear combination of all indicators from the appropriate model specification is used. This index is then re-scaled to a [0,100] interval to allow a comparative index across gender. The new re-scaled HI, $HI^{[0,100]}$, is derived as follows:

$$HI_{it}^{[0,100]} = \frac{HI_{it} - HI_{min}}{HI_{max} - HI_{min}} \quad (15)$$

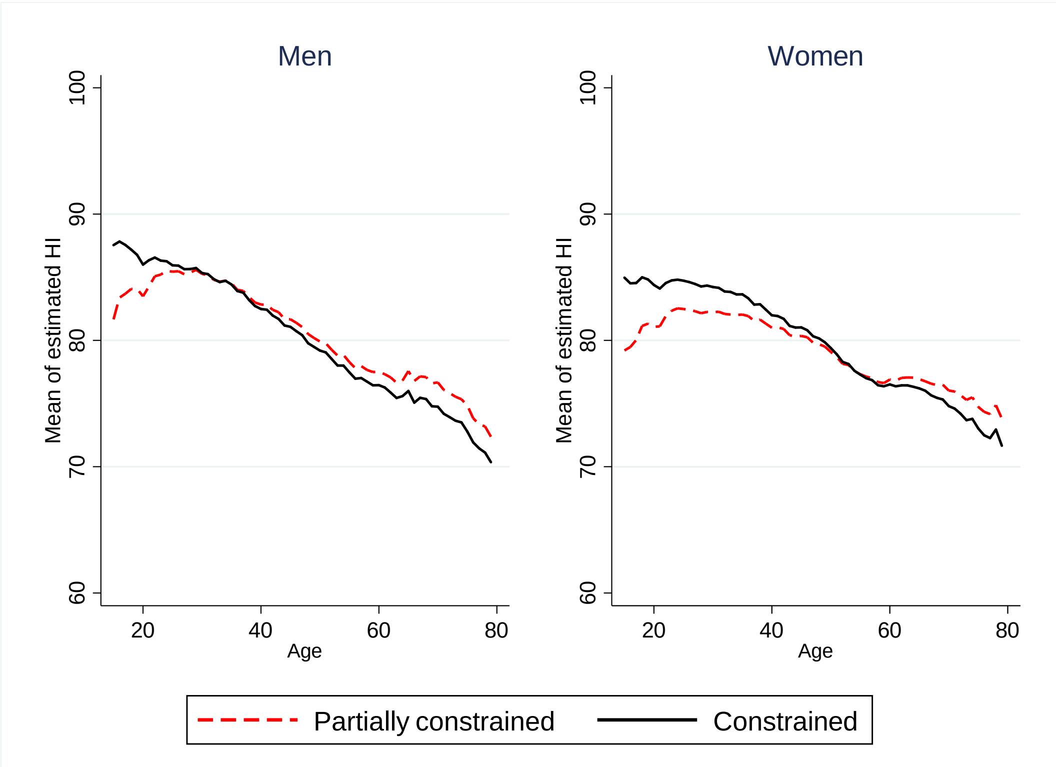
where HI_{it} presents the initially predicted linear HI for each individual i at time t, HI_{min} is the smallest and HI_{max} is the largest linear prediction across both all individuals i and time t. The standardisation of the estimated HI is similar to the approach employed by van Doorslaer and Jones (2003) and Pfarr *et al.* (2012a, 2012b).

Figure 1 shows mean values of the standardised estimated HI by age, gender and estimation technique and reveals gender differences with regards to health and reporting behaviour. Health is very similar for men and women in their sixties and seventies but men are in better health when they are younger. Although it is not surprising that individuals are more optimistic about their health when they are younger, the difference between the solid (constrained) and the dashed line (partially constrained) reflects reporting bias and therefore shows that the standard ordered probit model overestimates men’s health up until the early

[9] We use Williams’ (2006) user written “gologit2” command with the “autofit” extension to the GOP regression in Stata. It essentially starts out with a fully unconstrained model and then performs Wald tests on each variable to test whether it differs significantly across equations.

thirties and women's health much later when they are in their fifties, and underestimates health after those ages. People tend to be overly optimistic about their health when they are relatively young and report overly pessimistic health as they age.

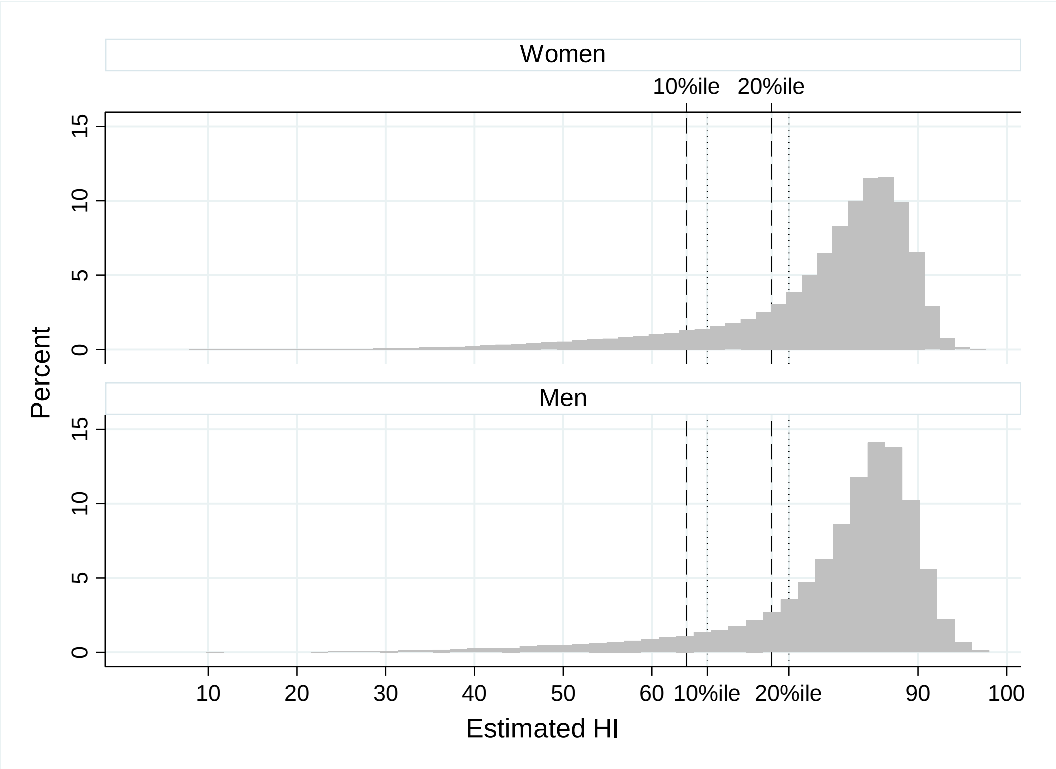
Figure 1. Mean Health Index by Estimation Technique, Gender and Age



Source: UKHLS and authors' own calculations

To address '*Type II endogeneity'* caused by potential reporting or error-in-variable bias, the estimated HI, rather than SAH or purely objective health measures, is employed both as a continuous and a binary variable. For the latter, we allocate respondents into poor relative to good health if they are in the bottom 10 percentile, or fair relative to good health if they are in the bottom 20 percentile, respectively, of the HI distribution, illustrated in Figure 2.

Figure 2. Distribution and Defined Cut-off Points of the Estimated Health Index



Source: UKHLS and authors' own calculations

It shows that most of the population is in relatively good health with a long tail on the left and points where the distributions increase sharply. The bottom 10 percentile is in line with the amount of people, who are self-assessing their health as poor and it is still at a point on the distribution where there is no significant increase to be observed. The bottom 20 percentile is assumed to reflect poor to fair health because the distribution jumps for both men and women at this point. The use of a binary variable allows an easier interpretation, especially when interacted with spouses´ characteristics, and these binary variables are used in the HCS estimation[10].

[10] The binary construct of the HI is used to identify the number of people in poor or fair health. Due to relatively small cell sizes, i.e. number of people in poor or fair health by year, gender, qualification and age, and the short time period available in the UKHLS, we take the average number of people in poor or fair health across all available waves and take that as the midpoint in our time series. With the help of health variables taken from the LFS, we construct a time trend that is then applied to that midpoint to extend the time period of our analysis from 1996 to 2018.

3.3 Health and Labour Supply

This sub-section reports on our attempt to estimate the effect of health level changes and health shocks on retirement for individuals aged up to 70 using random effect probit regressions. Since retirement decisions are often made jointly between spouses (Jones *et al.* 2010), the study also accounts for their health as well as their employment status. Furthermore, we control for caring activities within as well as outside of their own household to incorporate individual, spouse's and third parties' health (Dentinger and Clarkberg, 2002; van Houtven *et al.* 2013).

The relationship between health and retirement is endogenous because individuals have to make choices in the production of health capital jointly with labour supply and other consumption (Grossman, 1972). Economic activity can have a direct effect on health if either non-participation or participation leads to a deterioration of health through boredom or work pressure and bad working conditions, respectively. This reverse effect of economic status and health introduces '*simultaneity bias'* if health is treated as exogenous in retirement models. To address this '*type I endogeneity'*, health dynamics are exploited, which is a common strategy in the literature (see Bound *et al.* 1999; Disney *et al.* 2006; Hagan *et al.* 2008; Jones *et al.* 2010; García-Gómez *et al.* 2010). Therefore, we condition past health, $\widehat{HI_{it-1}}$, on initial health, $\widehat{HI_{it=0}}$, to represent health shocks. By using time variations in health, unobserved heterogeneity is minimised and concerns for reverse causality between health and retirement are reduced. Standard in the literature, we test for potential '*selection bias'* into employment by applying Heckman's (1976, 2013) two-step estimation procedure. We use random effect probit regressions and regress retirement status, $retired_{it}$, of an individual *i* at time *t* on the previously estimated HI, spouse's characteristics, caregiving responsibilities and a vector of control variables, X_{it}[11]. Spouse's characteristics capture if the spouse was in poor health based on the estimated HI, $HI\widehat{10_{SP}}_{it-1}$, and/or retired, $retiredSP_{it-1}$, in the previous year. To analyse the effect of caregiving duties on retirement, dummy variables are included to indicate whether the respondent is looking after someone sick, handicapped or elderly within ($carerHH_{it-1}$) or outside of their household ($carer\_xHH_{it-1}$,) or both. Since there is

[11] They include age, age squared, highest educational or vocational qualification, regional dummies, the initial employment sector at the first wave of observation and household wealth, such as household income and housing tenure.

evidence that employment participation and earnings affect the willingness to provide informal care negatively (see Carmichael *et al.* 2010), the caregiving variables are lagged. The full specification is given in equation (16).

$$\Pr\left(retired_{it} = 1 \middle| \widehat{HI_{it=0}}, \widehat{HI_{it-1}}, spouse_{it-1}, \widehat{HI\_SP}_{it-1}, retiredSP_{it-1}, carer_{it-1}, X_{it}\right) = \beta_1 \widehat{HI_{it=0}} + \beta_2 \widehat{HI_{it-1}} + S_{it-1} + C_{it-1} + \delta' X_{it} + \varepsilon_{it} \quad (16)$$

where $retired_{it}$ = 1 if $ret_{it} > 0$ and 0 otherwise

S_{it-1}= $\beta_2 spouse_{it-1} + \beta_3 spouse_{it-1} HI\widehat{10_{SP}}_{it-1} + \beta_4 spouse_{it-1} retiredSP_{it-1} + \beta_5 spouse_{it-1} HI\widehat{10_{SP}}_{it-1} retiredSP_{it-1}$,

C_{it-1}= $\beta_6 carerHH_{it-1} + \beta_7 carer\_xHH_{it-1} + \beta_8 carerHH_{it-1} carer\_xHH_{it-1}$

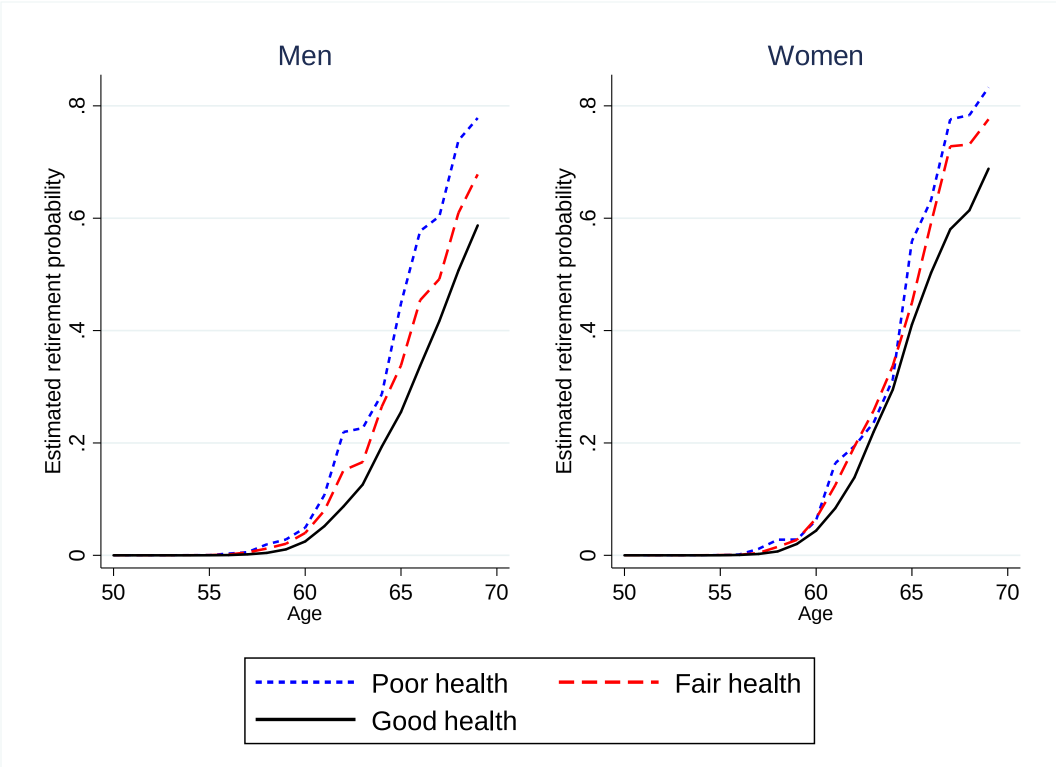
ε_{it}= time varying error term, which is uncorrelated with the other regressors

We find health to be a highly significant determinant of men's and women's retirement decision[12]. While the health effect is very similar in magnitude for both men and women (see Appendix A3 for the regression results), Figure 3, which illustrates graphically the estimated retirement probabilities by gender and health status, highlights gender differences. Therefore, men in poor health are 2.2 times and in fair health 1.8 times more likely to retire than men in good health. These ratios are 1.7 and 1.4 for women, respectively. Furthermore, our results support Jones *et al.*'s (2010) findings that couples are likely to make joint retirement decisions. This effect is much larger for women, indicating that their spouses retire first and women follow, and also applies if their retired spouse is in poor health (this is not the case for men's spouses). However, unlike for the previous literature, we do not observe any significant health effects of third parties on retirement[13].

[12] The effect is bigger for contemporaneous health rather than past health, which can be a result of the potential endogeneity bias, introduced through reverse causality, or because retirement decisions are made immediately after the onset of health deterioration.

[13] Ideally, we would know the relationship between the caregiver and the recipient, rather than just the residential circumstances, since it is likely to affect the extent of care provided and the labour response. If the recipient and the caregiver are family members, the endogenous nature of caregiving could be addressed by using the number of siblings of the caregiver as an instrument.

Figure 3. Comparison of Estimated Retirement Probabilities by Health, Gender and Age



Source: UKHLS and authors' own calculations

We also applied Heckman's two stage procedure to two different samples in the outcome model using the number of children below the age of 16 as the exclusion restriction[14]. To take into account possible selection bias when constructing health adjusted retirement probabilities for the HCS estimation, consecutive waves of employment history are required which only allow for transitions into retirement. However, when estimating the probability of being employed in the selection model, the sample has to include individuals inside and outside of the labour force, which leaves gaps in the following retirement model if individuals move out of employment and into unemployment, family care, etc. We call this the sample with the 'discontinuous' employment spell. In addition, we also test for selection bias by dropping these observations in the outcome equation, so that a sample with only 'continuous' employment spells remains. Although this will lead to a comparison of two different

[14] We argue that dependent children are correlated with the selection into employment but not with the decision to retire because they are likely to incentivise male and discourage female labour participation. This diminishes with age and, thus, becomes less relevant in our retirement model.

population groups when examining the effect of the Inverse Mills ratio, it is the more appropriate sample for the retirement model. Although selection bias is only evident in the continuous employment spell sample, this specification is conceptually not satisfactory to test for selection bias because we are essentially looking at different groups of people. By dropping additional observations between the two stages, we predict employment behaviour on the whole sample but then only predict retirement behaviour on respondents, who stay in employment until they retire. When we predict for selection into employment, the coefficient on health is larger in absolute value for both genders and across both specifications, which indicates that the health effect on retirement is underestimated. However, the difference is only large in the conceptually unsatisfactory continuous employment spell model. Therefore, in the HCS estimation we employ the more conservative health estimate and do not account for selection bias.

Although we do not estimate health effects on hours, we know how many hours a person in good and a person in poor health works. On average, men and women in fair health work 4 per cent and 5 per cent less than healthy people, respectively. These numbers increase to 6 and to 7 per cent when for men and women in poor health, respectively.

3.4 Health and wages

Although the literature is replete with clever natural experiments designs in attempts to link health to wages using data on one off events, such as accidents, we focus on the literature which tries to find a relationship using more general SAH based measures. This literature presents evidence that good health increases men's wages significantly relative to poor health, with weaker effects on female wages (see e.g. Contoyannis and Rice, 2001; Cai, 2009; Jäckle and Himmler, 2010; Flores and Kalwij, 2013, with the evidence in Brown *et al.* 2010, showing effects also for men).

However, the relationship is associated with econometric challenges by introducing different sources of potential endogeneity. Firstly, as was the case in the previous two sections, the health measures employed are often associated with measurement error (type II endogeneity). Therefore, we continue to use the previously estimated HI to provide information on the severity as well as on the prevalence of certain health conditions, address heterogeneous reporting behaviour and reduce concerns for justification bias. Secondly, endogeneity is likely to be introduced by the correlation between unobserved factors, such as genetic endowment, which impact on health and wages at the same time. This issue of unobserved heterogeneity is of similar source as the one discussed in the previous section and, hence, it is addressed in the same way. We exploit the variation of health over time by employing a health shock. Thirdly, there is a reverse effect of wages on health and it is theoretically unclear whether this is positive or negative. The health production model by Grossman (1972) views health as an endogenously determined capital stock, which depreciates with age, appreciates with investments and changes with allocated time between work and leisure. With greater capital stock, more time becomes available to spend on productive activities and utility increases. This implies that higher wages increase the marginal return to health proportionally and increase opportunity costs of health investments. Therefore, less time is invested in the production of health and more is invested in the labour market (Grossman and Benham, 1974), which suggests a negative effect of wages on health. However, the same health production model also argues that investments in health capital, to maintain or improve its stock, are constrained by the availability of time and economic resources. Therefore, higher wages can also improve health by increasing economic inputs available for the production of health. We reduce concerns for reverse causality by lagging the HI by one period. Moreover, we estimate health effects on wages by qualification. It is assumed that better health increases productivity and higher wages make better health more affordable through better treatments and nutrition. Since higher income is associated with higher qualifications, it can be argued that running regressions by qualification groups reduces, but does not entirely eliminate, this source of endogeneity. The effect of health on wages is measured using a Mincerian type wage equation (Mincer, 1974), as shown in equation 17, and is run by gender and estimated using OLS[15].

$$lnwage_{it} = \beta_0 + \beta_1 \widehat{HI_{it=0}} + \beta_2 \widehat{HI_{it-1}} + S_{it-1} + C_{it-1} + \gamma' X_{it} + \varepsilon_{it} \qquad (17)$$

According to Grossman (1972), health impacts on productivity and, as a result, can be considered an endogenous capital stock, which affects the amount of hours spent participating in the labour force. Since health differences can determine working hours, we

[15] We do not apply any fixed effect specification because it omits the time-invariant initial HI observed in the first period, which is part of the health shock measure.

use log hourly wages, $lnwage$, rather than monthly or annual earnings, when measuring the health effect on wages. Since there is no hourly wage variable in the UKHLS, it is constructed using data on gross monthly pay and hours worked, similar to that used by Contoyannis and Rice´s (2001), and uses income from respondents' main and their potential second job. If individuals have relatively low wages, they are more likely to subsidise their income with another job.

Similar to the retirement model, the wage equation also contains information on cohabitating couples and caregiving duties. If cohabitating couples benefit from household economies of scale within a household production framework, more of their time can be allocated towards labour market activities. Equally, spouse's retirement status and their health are likely to influence the allocation of work in the household production function. The responsibility associated with caregiving commitment reflects time constraints and, hence, it presents a trade-off between the time spent on caregiving and the time spent on labour market activities.

In the control variables, X, we include information on occupational status, part-time employment, employer's firm size and employment sector, since previous research for the UK observed wage differentials between the private and public sector (e.g. Disney and Gosling, 1998), especially for women. The wage equation also controls for a quadratic function of age and work experience in the last job. While age is assumed to capture general market experience implicitly, (Mincer, 1974), work experience, conditioned on age, captures job-specific knowledge and training obtained at the current job. Since Mincer's model assumes that the amount of time spent on training declines over the life-cycle, concave age and experience functions are expected. We also include information on the highest educational or vocational qualification obtained by the respondent, regional dummies and a vector of wave dummies to control for aggregate productivity effects and inflation.

Wages can only be observed for working individuals. Since individuals in very poor health are unlikely to work, the true effect of health on wages is not captured if only economically active respondents are studied. We control for non-random self-selection into work (for self-employed respondents and employees) using Heckman's two-stage procedure (1976) and employing house ownership and number of young children as restriction terms for men and women, respectively. Much of the literature relies on non-wage income, in particular on spouse's income, which is observed to be negatively correlated with respondents' labour income. It can be argued that above a certain threshold, alternative income sources incentivise economic inactivity. However, exclusion restrictions are only valid selection instruments if they neither affect respondents' wages directly, nor are they correlated with unobserved factors affecting wages. Since the existing literature suggests that high achieving individuals with superior productivity tend to marry one another (Becker, 1981) and have similar earnings even prior to their marriage (Nakosteen *et al.* 2004), the latter assumption is likely to be violated. There appears to be unobserved characteristics, which are related to expected earnings and the probability of finding a partner with a specific earning potential. As a result, spouse's labour income and household income net respondent's labour income have been rejected as potential exclusion restrictions. Although household income net of all labour income can be a suitable alternative, it is still likely to be correlated with some unobserved attributes, such as ability or motivation, affecting wages at the same time. Consequently, we use housing tenure as an exclusion restriction. It is likely to be correlated with the selection into work (or rather the selection out of work) because it reflects financial accomplishment. While outstanding mortgages or monthly rent payments are a financial long-term commitment, outright homeowners do not have this financial burden any longer. Furthermore, while wages affect house ownership directly, there is less concern for any reverse causality.

For women, we use the number of children below the age of two living in the household. While it is negatively correlated with women's labour supply because of the time constraint associated with child caring, our previous findings on labour supply suggest that children have no significant effect on men's decision to work. This does not change when the children's age is reduced and is also supported by the literature, which identifies significant negative effects of fertility on labour supply among mostly women. Although cohabitation with young children is not necessarily associated with child rearing, the young age range in this variable suggests that child caring affects all household members one way or another. The USS provides information on the number of children the respondent is responsible for, which can be argued to be more suitable. However, children are between the ages of 5 to 16 in this variable and, consequently, are less dependent and less likely to affect the respondent's employment decision. Although there are theoretical arguments that fertility, labour supply and wages are endogenous when low wages are correlated with higher fertility, it has been identified as a valid selection instrument by numerous scholars (see Mulligan and Rubinstein, 2008; Huber and Mellace, 2014). Moreover, estimated health effects are almost identical when employing house ownership or young children as exclusion restrictions in the female wage equation.

Our results (see Appendix A4 and A5 for the selection and the wage equation, respectively) show that good health is positively and significantly associated with hourly wages with a slightly higher health effect among men. After addressing all econometric concerns, our results show that wages increase by up to 1.2 per cent and up to 1.0 per cent for men and women, respectively, with every one-unit increase in the estimated HI. Although this is at the lower end of the spectrum compared to existing work on health and wages, the nature of our health measure makes comparisons difficult. Therefore, Table 1 shows predicted hourly wages by gender and health status using the constructed dummy variables indicating poor and fair health discussed above. SAH, which is provided by the UKHLS, is also included for comparative purposes. We find poor health to decrease hourly wages for men and women by 21 per cent and 16 per cent, respectively. These wage differentials are reduced to 16 per cent and 11 per cent when fair and good health are compared. The size of these effects are in line with Cai's (2009) findings, who observed a wage differential of 20 per cent when using predicted health categories based on his HI and comparing excellent or good health to poor or fair health.

Table 1. Predicted Hourly Wages from OLS Model with Exclusion Restrictions

Constructed health dummies based on the estimated HI	Men	Women
Poor vs. good		
Good	13.77	11.16
Poor	10.82	9.33
Change	-21%	-16%
Fair vs. good		
Good	13.88	11.21
Fair	11.69	9.96
Change	-16%	-11%
SAH		
Excellent/very good/good	13.88	11.25
Poor/fair*	12.20	10.10
Change (excellent to poor/fair)	-12%	-10%

*Source: UKHLS and authors' own calculations; *equivalent to Fair in the 2nd panel.*

4. Health and Human Capital

This section brings together the estimates discussed in section 3 with the theoretical framework in section 2. We first discuss the basic data setup and then present some results.

4.1 Data for Human Capital Stocks

The database is constructed in a similar way to the approach adopted by the Office for National Statistics (ONS) in their national annual human capital accounts and therefore relies mainly on data from the LFS. Information is collected on the number of people aged 16 to 69 (i.e. the working age population), their annual earnings, enrolment probabilities for schools, further and higher education as well as employment and survival rates[16] for the years 1996 to 2018. A detailed description of all variables used and their sources is provided in Appendix A6. Compared to the estimations provided by the ONS, we use information beyond the age of 64 and also collect data on unemployment, retirement and sickness rates in order to incorporate the estimated health effect on labour supply in terms of changes in retirement behaviours. With the exception of the survival rate, which only varies by gender and age, all

[16] Life tables and the longitudinal version of the LFS are currently only published until 2017. For now, we assume the same survival rates and educational enrolment probabilities in 2017 also apply for 2018. However, ONS will release the next set of life tables in September 2019.

other information is cross-classified by gender, age, qualification and health status. Qualifications are classified into the following six categories, which allow the coherent use of several datasets:

- No qualification or don't know
- Other qualification
- GCSE or equivalent
- A-levels or equivalent
- Further education
- Degree and higher or equivalent

Some imputation and adjustments are necessary to address missing observations, especially at the lower and upper tails of the age distribution[17] but these are carried out in such a way to ensure compatibility with ONS estimates[18].

The estimated wage differential in lifetime labour earnings is incorporated by transforming health adjusted hourly wages from the UKHLS into annual earnings by age, gender and qualification individually. Since the LFS is used for the overall HCS estimation (see below), earnings from the LFS are weighted by these estimates to arrive at health adjusted annual earnings.

4.2 Results

This section shows the results of combining the impacts of health on earnings and retirement with the JF HCS model. First, it reviews detailed findings on the population. Later we show the difference between potential and productive HCS and how much of this is due to additional retirement probability of those in poor health and those inactive due to long-term illness. For presentation purposes we only report the findings obtained when incorporating poor (the bottom 10 percentile of the HI) rather than fair health (the bottom 20 percentile). The latter is presented in Appendix A7 to A10.

---

[17] For instance, we assume earnings are zero for degree holders aged 16 to 18 and further education holders aged 16.

[18] For instance, population numbers are tabulated from the LFS by gender, age and qualification using population weight and benchmarked to population numbers taken from ONS, which are only provided by gender and age. Any difference in population number between both totals is divided equally across all qualification groups and added/subtracted to arrive at the total population number provided by the LFS

Figure 4a shows the distribution of the population aged 16-69 by activity and health status in 2018 where good health is distinguished from poor health using the bottom 10 percentile of the HI (excluding those enrolled in education), and Figure 4b shows the time series information. Appendix A7 displays the same pie charts using fair health. The majority of this population are in employment and in good health, with a greater proportion of males than females in this category, and a corresponding difference in the 'Other' category that includes inactive due to child and adult caring responsibilities. A small proportion of those employed are in poor health. There is also a sizable category who are inactive due to long-term illness, and a high proportion of retirees are in poor health. Over time, there is some reduction in the proportion active due to long-term illness, most noticeably for males, but other categories are relatively stable.

Figure 4. Population Aged 16-69 by Job and Health Status (in per cent)

Figure 4a. 2018

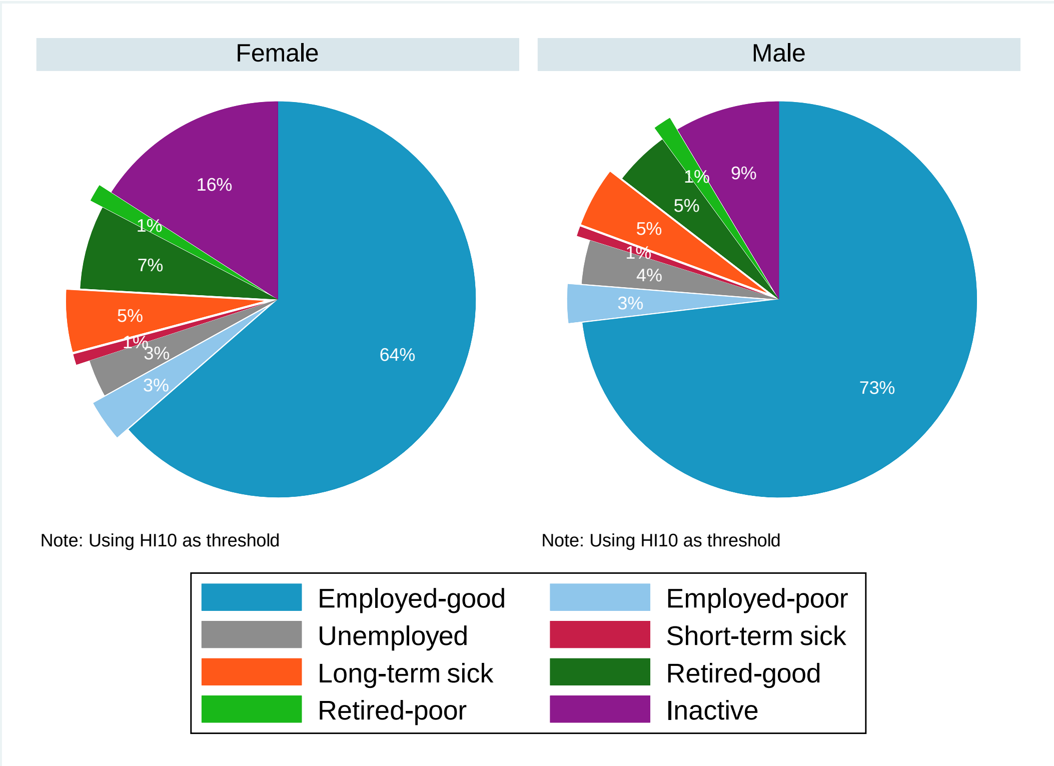
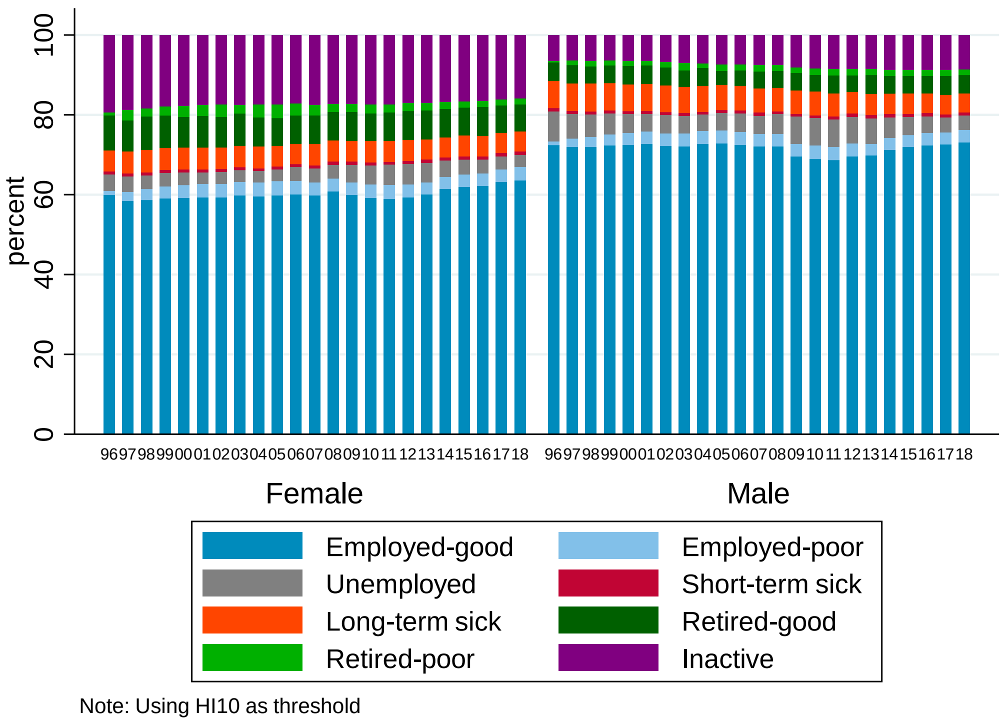


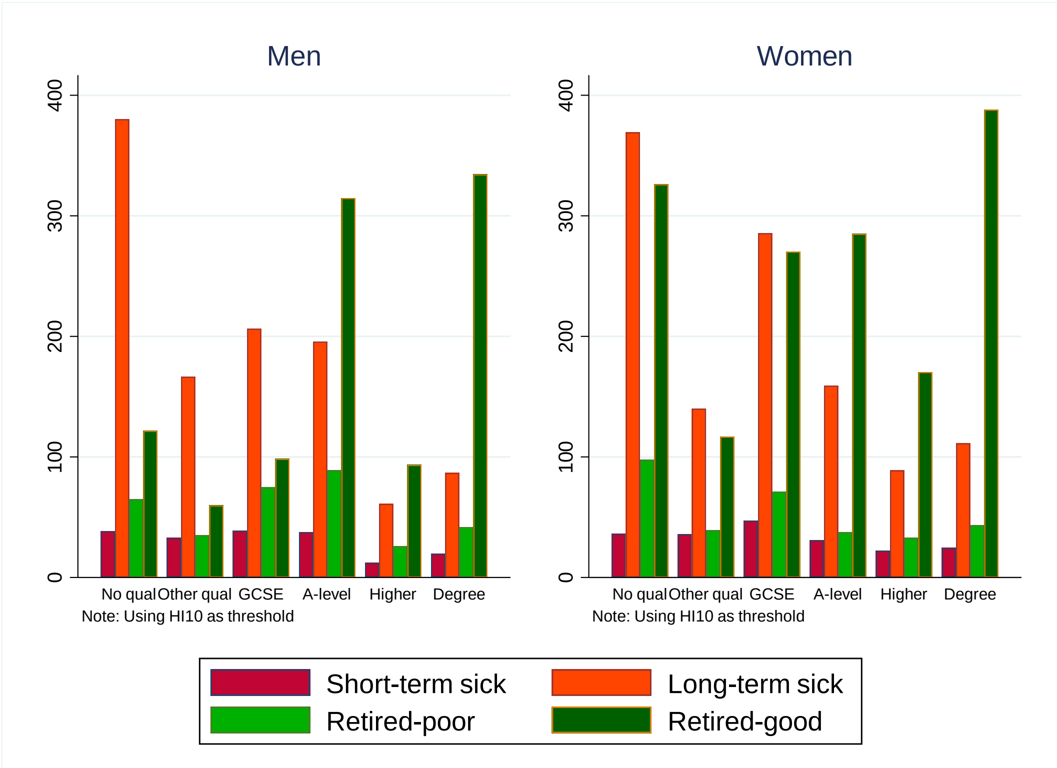
Figure 4b. 1996-2018



Source: UKHLS, LFS and authors' own calculations

Figure 5 shows a snapshot of qualifications and health status for the same population in 2018. It shows clearly that long-term illness and retirement due to poor health is concentrated in the lower qualification categories. The equivalent bar charts using fair health are presented in Appendix A8.

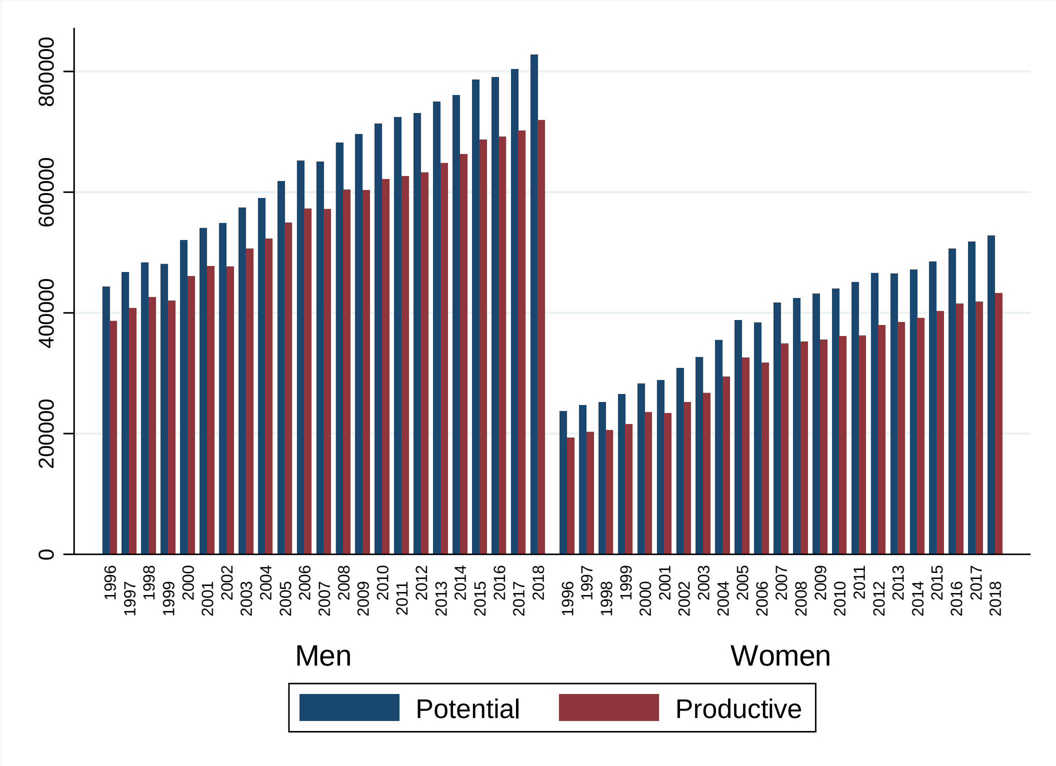
Figure 5. Population Aged 16-69 by Qualification and Health Status (in thousand), 2018



Source: UKHLS, LFS and authors' own calculations

Figure 6 shows the average potential and productive HCS per person for both men and women in nominal terms for the time period 1996 to 2018. Both, potential and productive average HCS have been increasing over time but the gap has also been rising. Although there was some catching up of females to males, the gender differential is still pronounced by the end of the period.

Figure 6. Nominal Potential and Productive Average Human Capital Stock by Gender (in GBP), 1996-2018



Source: UKHLS, LFS and authors' own calculations

Table 2 shows the ratio of potential to productive HCS and how much of this is due to absence and early retirement due to sickness, in total and for men and women over the time period 1996 to 2018. It defines ill health using the bottom 10 percentile of the HI. The equivalent results when using the bottom 20 percentile of the distribution are presented in Appendix A9. It shows that overall potential HCS is about 60% higher than productive HCS, throughout the period and about 20% of this difference is due to poor health. The health effect is mostly driven by long-term illness. This percentage is relatively stable over time. The potential HCS is much higher than the productive HCS for women than for men, due to greater inactivity rates by women. We also estimate that the productive HCS is about 1% lower than it would be if all people were in good health, due to lower annual earnings of those in poor health, and this is similar across gender and time. This relatively small impact of lower earnings reflects the fact that only a small proportion of those in employment are in poor health.

Table 2. Sickness, Presenteeism and Absenteeism, 1996-2018

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Total																							
Productive HCS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Potential HCS	**1.65**	**1.62**	**1.60**	**1.60**	**1.57**	**1.57**	**1.59**	**1.57**	**1.56**	**1.55**	**1.56**	**1.56**	**1.55**	**1.61**	**1.61**	**1.63**	**1.63**	**1.61**	**1.58**	**1.56**	**1.56**	**1.55**	**1.54**
Difference due to (%)																							
Long-term ill	17.7	18.7	19.7	19.0	19.5	19.3	18.9	18.6	19.2	17.9	16.9	17.4	17.4	15.8	15.6	14.9	14.4	14.6	15.1	15.5	15.4	15.6	16.2
Short-term ill	2.9	2.9	2.8	2.7	3.0	2.7	2.7	2.9	2.7	2.7	2.9	3.0	2.9	2.7	2.7	2.7	2.8	3.0	3.0	3.1	3.1	3.2	3.2
Earnings (ill-health)	0.4	1.0	1.3	1.4	1.6	1.6	1.6	1.7	1.7	1.7	1.7	1.7	1.6	1.5	1.7	1.6	1.6	1.5	1.5	1.6	1.7	1.7	1.8
Retirement (ill-health)	0.2	0.6	0.7	0.9	0.9	1.0	1.1	1.0	1.1	1.1	1.1	1.0	1.0	1.0	1.1	1.0	1.1	1.0	1.0	1.0	1.1	1.2	1.2
Retirement (general)	20.5	22.8	23.7	24.9	25.7	26.9	27.3	27.2	27.6	27.8	28.3	26.8	26.5	25.2	24.6	23.8	25.2	26.0	27.2	28.1	28.1	29.7	29.5
Other	58.2	54.0	51.9	51.1	49.2	48.4	48.5	48.6	47.8	48.8	49.0	50.1	50.6	53.8	54.3	56.0	54.9	54.0	52.2	50.7	50.6	48.7	48.1
Men																							
Productive HCS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Potential HCS	**1.51**	**1.49**	**1.46**	**1.47**	**1.44**	**1.44**	**1.46**	**1.44**	**1.42**	**1.42**	**1.44**	**1.43**	**1.42**	**1.49**	**1.49**	**1.50**	**1.50**	**1.50**	**1.47**	**1.45**	**1.44**	**1.44**	**1.44**
Difference due to (%)																							
Long-term ill	23.4	24.7	26.5	25.7	26.5	25.6	25.2	25.1	25.8	24.4	22.6	23.0	23.1	20.4	20.1	19.4	18.3	18.4	18.9	18.8	19.2	19.5	19.7
Short-term ill	3.5	3.4	3.3	3.1	3.5	3.2	3.3	3.6	3.5	3.4	3.7	3.7	3.5	3.1	3.2	3.1	3.4	3.5	3.5	3.7	3.8	3.8	3.8
Earnings (ill-health)	0.6	1.3	1.7	1.9	2.1	2.2	2.2	2.3	2.3	2.4	2.2	2.3	2.3	2.0	2.3	2.2	2.1	2.0	2.0	2.1	2.3	2.3	2.4
Retirement (ill-health)	0.2	0.5	0.7	0.8	0.9	1.0	1.1	1.0	1.0	1.1	1.1	1.1	1.0	0.9	1.1	1.0	1.1	1.1	1.1	1.2	1.2	1.2	1.3
Retirement (general)	20.1	24.0	25.1	25.9	27.2	28.7	29.3	27.8	28.6	28.9	31.0	29.7	28.1	25.9	26.0	25.6	27.4	28.9	30.7	32.4	31.8	32.9	33.4
Other	52.1	46.0	42.7	42.6	39.8	39.4	39.0	40.2	38.7	39.9	39.3	40.2	42.1	47.7	47.3	48.6	47.6	46.0	43.7	41.9	41.7	40.3	39.5
Women																							
Productive HCS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Potential HCS	**1.97**	**1.94**	**1.93**	**1.92**	**1.86**	**1.90**	**1.87**	**1.86**	**1.83**	**1.79**	**1.82**	**1.80**	**1.81**	**1.84**	**1.85**	**1.89**	**1.87**	**1.82**	**1.79**	**1.77**	**1.78**	**1.78**	**1.74**
Difference due to (%)																							
Long-term ill	10.5	11.1	11.2	11.1	11.2	11.8	11.3	11.3	12.0	11.1	10.6	11.6	11.4	10.6	10.5	9.8	10.1	10.2	10.8	11.7	11.3	11.5	12.4
Short-term ill	2.3	2.3	2.1	2.1	2.4	2.0	2.1	2.1	1.9	1.9	2.0	2.2	2.4	2.2	2.1	2.2	2.1	2.3	2.3	2.5	2.4	2.5	2.5
Earnings (ill-health)	0.2	0.6	0.8	0.8	0.9	1.0	0.9	1.0	0.9	1.0	1.0	1.0	0.9	1.0	1.0	1.0	1.0	0.9	0.9	1.0	1.0	1.1	1.2
Retirement (ill-health)	0.2	0.6	0.8	0.9	1.0	1.1	1.0	1.1	1.1	1.1	1.1	1.0	1.0	1.0	1.1	1.0	1.0	0.9	0.9	0.9	1.0	1.1	1.1
Retirement (general)	20.8	21.4	22.0	23.8	24.0	24.8	24.9	26.5	26.5	26.7	25.4	23.8	24.8	24.3	23.0	21.8	22.8	22.5	23.3	23.3	24.0	26.2	25.2
Other	65.9	64.1	63.2	61.2	60.5	59.2	59.7	58.1	57.6	58.1	59.9	60.5	59.6	60.8	62.3	64.2	63.0	63.2	61.8	60.7	60.3	57.6	57.6

Source: UKHLS, LFS and authors' own calculations

A somewhat different pattern emerges if we further distinguish between qualification (low and high) and age (30 and years and older as well as 50 years and older). Table 3 presents the ratios of potential to productive HCS by age and qualification level in 2018 for the health effects. The equivalent results when using the bottom 20 percentile of the distribution are presented in Appendix A10. The effects are much larger for those with low qualifications and especially for the group aged 50 and over. The latter group are also most impacted by the health effects of earnings - the productive HCS was nearly 4% lower for this group relative to the counterfactual of all persons being in good health.

Table 3. Sickness, Presenteeism and Absenteeism by Gender, Qualification and Age Group, 2018

	Low qualification			**High qualification**		
In 2018	**F**	**M**	**T**	**F**	**M**	**T**
			Aged 30+			
Potential HCS	**2.34**	**1.59**	**1.81**	**1.60**	**1.35**	**1.45**
Long-term ill	1.21	1.16	1.17	1.05	1.04	1.04
Short-term ill	1.03	1.03	1.03	1.01	1.01	1.01
Earnings (ill-health)	1.03	1.03	1.03	1.01	1.01	1.01
Retirement (ill-health)	1.04	1.02	1.02	1.01	1.00	1.00
Retirement (general)	1.31	1.18	1.22	1.20	1.17	1.19
			Aged 50+			
Potential HCS	**2.88**	**2.14**	**2.40**	**2.24**	**1.90**	**2.02**
Long-term ill	1.31	1.28	1.29	1.11	1.07	1.09
Short-term ill	1.04	1.03	1.04	1.02	1.01	1.01
Earnings (ill-health)	1.03	1.04	1.04	1.01	1.02	1.01
Retirement (ill-health)	1.14	1.09	1.11	1.05	1.04	1.04
Retirement (general)	1.76	1.45	1.56	1.66	1.54	1.58

Source: UKHLS, LFS and authors' own calculations

5. Conclusion

This paper shows the importance of incorporating health in HCS estimations by highlighting the significant impact of health on human capital through its effect on inactivity, early retirement and wages. This work reveals that health effects can be studied in detail with different thresholds for health statuses in mind, which allows changes in HCS to be quantified in different contexts. In general, self-assessed health overestimates those in poor and fair health at the expense of those in good and very good health. On average, individuals with poor health are about twice as likely to retire than people in good health, with a higher ratio for males than females, and couples tend to synchronise their retirement. On average, women in poor health earn about 16% lower hourly wages of those in good health, whereas for men the figure is 21 per cent. When we combine these estimates with the lower activity rates due to illness within a lifetime income framework, we can estimate the impact on potential Human Capital Stocks. We show that overall poor health leads to a reduction in HCS which is mostly due to long-term illness. The impact of presenteeism is relatively small, due to the low percentage of the working population in poor health. This effect is larger for those with low qualifications, especially among the 50+ age group. Absenteeism due to permanent retirement is also very small overall, but this channel only applies to the over 50 age group and for the latter poor health has a very significant impact on the decision to retire.

The low values overall for presenteeism is perhaps the most surprising finding as this is a frequent topic of discussion in policy circles. However, these effects do not take account of any spillovers from workers in poor health to those in good health. Further investigation of the aggregate impact of such spillovers is necessary, (Kinman, 2019), since existing research of the impact of presenteeism on others is mainly carried out in the health care environment (Forsythe *et al.* 1999; Halbesleben and Rathert, 2008; Demerouti *et al.* 2009; Letvak *et al.* 2012).

References

Ahlroth, S., Bjorklund, A. and Forslund, A. (1997) The Output of the Swedish Education Sector. *Review of Income and Wealth,* 43 (1): 89–104

Anderson, K.H. and Burkhauser, R.V. (1985) The retirement-health nexus: a new measure of an old puzzle. *Journal of Human Resources*, 20 (3): 315–330

Atkinson, T. (2005) *Atkinson Review: Final Report: Measurement of Government Output and Productivity for the National Accounts*. Basingstoke: Palgrave MacMillan

Barro, R. J. (1996) *Health and economic growth*. Cambridge, MA: Harvard University

Becker, G. S. (1981) Altruism in the Family and Selfishness in the Market Place. *Economica*, 48 (189): 1-15

Becker, G. S. (1964) *Human Capital*. New York: Columbia University Press

Bound, J. (1991) Self-Reported vs. Objective Measures of Health in Retirement Models. *Journal of Human Resources*, 26 (1): 106-138

Bound, J., Schoenbaum, M., Stinebrickner, T. R. and Waidmann, T. (1999) The Dynamic Effects of Health on the Labor Force Transitions of Older Workers. *Labour Economics*, 6 (2): 179-202

Brown, S., Roberts, J., and Taylor, K. (2010) Reservation wages, labour market participation and health. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173 (3): 501-529

Burton, W. N., Conti, D. J., Chen, C., Schultz, A. B. and Edington, D. W. (1999) The role of health risk factors and disease on worker productivity. *Journal of Occupational and Environmental Medicine,* 41 (10): 863–877

Cai, L. (2009) Effects of health on wages of Australian men. *Economic Record*, 85 (3): 290–306

Carmichael, F., Charles, S. and Hulme, C. (2010) Who will care? Employment participation and willingness to supply informal care. *Journal of Health Economics*, 29 (1): 182-190

Christian, M. (2010) Human Capital Accounting in the United States, 1994–2006. *Survey of Current Business,* 90 (6): 31–6

Contoyannis, P. and Rice, N. (2001) The impact of health on wages: evidence from the British Household Panel Survey. *Empirical Economics*, 26 (4): 599-622

Dang, T., Antolin, P. and Oxley, H. (2001) *Fiscal Implications of Ageing: Projections of Age-related Spending* (No 305). Organisation for Economic Co-operation and Development

Demerouti, E., Le Blanc, P.M., Bakker, A.B., Schaufeli, W.B. and Hox, J. (2009) Present but sick: a three-wave study on job demands, presenteeism and burnout. *Career Development International*, 14 (1): 50-68

Dentinger, E. and Clarkberg, M. (2002) Informal caregiving and retirement timing among men and women gender and caregiving relationships in late midlife. *Journal of Family Issues*, 23 (7): 857-879

Disney, R. and Gosling, A. (1998) Does it pay to work in the public sector?. *Fiscal Studies*, 19 (4): 347-374

Disney, R., Emmerson, C. and Wakefield, M. (2006) Ill health and retirement in Britain: A panel data-based analysis. *Journal of Health Economics*, 25 (4): 621-649

Flores, M. and Kalwij, A. (2013) *What do wages add to the health-employment nexus? Evidence from older European workers.* Network for Studies on Pensions, Aging and Retirement Discussion Paper No. 03/2013-005

Forsythe M, Calnan M, Wall B. Doctors as patients: postal survey examining consultants and general practitioners' adherence to guidelines. Br Med J 1999; 319:605–8.

Fraumeni, B., Christian, M. and Samuels, J. (2017), The Accumulation of Human and Nonhuman Capital, Revisited. *Review of Income and Wealth*. 63 (S2), S381-S410.

García-Gómez P., Jones, A.M., Rice, N. (2010) Health effects on labour market exits and entries. *Labour Economics*, 17 (1): 62-76

Garrow V. (2016) Presenteeism: A Review of Current Thinking. *Institute for Employment Studies*, Available at: http://www.employment-studies.co.uk/system/files/resources/files/507_0.pdf (26 June 2019, date last accessed).

Goetzel, R. Z., Long, S. R., Ozminkowski, R. J., Hawkins, K., Wang, S. and Lynch, W. (2004) Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers. *Journal of Occupational and Environmental Medicine*, 46 (4): 398-412

Grossman M. and Benham, L. (1974) Health, hours and wages. In: Perlman, M. ed. *The Economics of Health and Medical Care*. New York: Halstead, pp. 205-233

Grossman, M. (1972) On the concept of health capital and the demand for health. *Journal of Political economy*, 80 (2): 223-255

Gu, W. and Wong, A. (2010) *Estimates of Human Capital in Canada: The Lifetime Income Approach* (Catalogue no. 11F0027M, no. 062). Economic Analysis Division, Statistics Canada

Hagan, R., Jones, A. M. and Rice, N. (2008) Health shocks and the hazard rate of early retirement in the ECHP. *Swiss Journal of Economics and Statistics*, 144 (3): 323-335

Halbesleben, J.R. and Rathert, C. (2008) Linking physician burnout and patient outcomes: exploring the dyadic relationship between physicians and patients. *Health care management review*, 33 (1):29-39

Hanusheck, E. A. and Kimko, D. D. (2000) Schooling, Labour force quality, and the Growth of

Nations. *The American Economic Review*, 90 (5): 1184-1208

Heckman, J. (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, 5 (4): 475–492

Heckman, J. (2013) Sample selection bias as a specification error. *Applied Econometrics*, 31 (3): 129-137

Hernández-Quevedo, C., Jones, A.M. and Rice, N. (2005) *Reporting Bias and Heterogeneity in Self-Assessed Health: Evidence from the British Household Panel Survey*. HEDG Working Paper 05/04, York: University of York

Huber, M. and Mellace, G. (2014) Testing exclusion restrictions and additive separability in sample selection models. *Empirical Economics*, 47 (1): 75-92

Idler, E. L. and Benyamini, Y. (1997) Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, 38 (1): 21-37

Idler, E. L. and Kasl, S. V. (1995) Self-ratings of health: do they also predict change in functional ability?. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 50 (6): 344-353

Jäckle, R. and Himmler, O. (2010) Health and Wages: Panel Data Estimates Considering Selection and Endogeneity. *The Journal of Human Resources*, 45 (2): 364-406

Johns, G. (2010) Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, 31 (4): 519-542

Jones, A. M., Rice, N. and Roberts, J. (2010) Sick of Work or too Sick to Work? Evidence on Self-Reported Health Shocks and Early Retirement from the BHPS. *Economic Modelling*, 27 (4): 866-880

Jones, R. and Fender, V. (2011) *Human Capital Estimates, 2010*. Office for National Statistics

Jones, R., and Chiripanhura, B. (2010) Measuring the U. K.'s Human Capital Stock. *Economic and Labor Market Review,* 4 (11): 36–63

Jorgenson, D. W. and Fraumeni, B. M. (1989) The accumulation of human and nonhuman capital, 1948–1984. In: Lipsey, R. E. and Tice, H. S. ed. *The Measurement of Savings, Investment and Wealth.* Chicago, I.L.: The University of Chicago Press, pp. 227–286

Jorgenson, D. W. and Fraumeni, B. M. (1992) The output of the education sector. In: Griliches, Z. ed. *Output Measurement in the Services Sector.* Chicago, I.L.: The University of Chicago Press, pp. 303–338

Kendrick, J. W. (1976) *The formation and stocks of total capital*. NBER Books.

Kinman, G. (2019) Sickness presenteeism at work: prevalence, costs and management. *British medical bulletin*.

Letvak, S.A., Ruhm, C.J. and Gupta, S.N. (2012) Nurses' presenteeism and its effects on self-reported quality of care and costs. *The American Journal of Nursing*, 112 (2): 30-38

Lindeboom, M. and Kerkhofs, M. (2009) Health and Work of the Elderly: Subjective Health Measures, Reporting Errors and Endogeneity in the Relationship between Health and Work. *Journal of Applied Econometrics*, 24 (6): 1024-1046

Loeppke, R., Hymel, P. A., Lofland, J. H., Pizzi, L. T., Konicki, D. L., Anstadt, G. W., Baase, C., Fortuna, J. and Scharf, T. (2003) Health-related workplace productivity measurement: general and migraine-specific recommendations from the ACOEM Expert Panel. *Journal of Occupational and Environmental Medicine*, 45 (4): 349-359

Lopez-Casanovas, G., Rivera, B. and Currais, L. (2005) *Health and Economic Growth*, Massachusetts: The MIT Press, pp.1-15

Mincer, J. (1974) Schooling, Experience, and Earnings. *Human Behavior & Social Institutions* No. 2.

Mulligan, C. B. and Rubinstein, Y. (2008) Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, 123 (3): 1061-1110

Nakosteen, R. A., Westerlund, O. and Zimmer, M. A. (2004) Marital matching and earnings evidence from the unmarried population in Sweden. *Journal of Human Resources*, 39 (4): 1033-1044

Pfarr, C., Schmid, A. and Schneider, U. (2012a) Reporting Heterogeneity in Self-Assessed Health Among Elderly Europeans. *Health Economics Review*, 2 (1): 1-14

Pfarr, C., Schneider, B. S., Schneider, U. and Ulrich, V. (2012b) I feel good! Gender differences and reporting heterogeneity in self-assessed health. *The European Journal of Health Economics*, 13 (3): 251-265

Samek, L. (2017) *Health, Labour Supply and Human Capital*, PhD thesis, King's College London.

Shultz, T. W. (1961) Investment in human capital. *American Economic Review*, 51 (1): 1-17

Stern, S. (1989) Measuring the effect of disability on labour force participation. *Journal of Human Resources*, 24 (3): 361–395

Turpin, R.S., Ozminkowski, R.J., Sharda, C.E., Collins, J.J., Berger, M.L., Billotti, G.M., Baase, C.M., Olson, M.J. and Nicholson, S. (2004) Reliability and validity of the Stanford Presenteeism Scale. *Journal of Occupational and Environmental Medicine*, 46 (11): 1123-1133

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 (9): 1621-1629

Van Doorslaer, E. and Jones, A. M. (2003) Inequalities in Self-Reported Health: Validation of a new Approach to Measurement. *Journal of Health Economics*, 22 (1): 61-87

Van Houtven, C. H., Coe, N. B. and Skira, M. M. (2013) The effect of informal care on work and wages. *Journal of Health Economics*, 32 (1): 240-252

Wei, H. (2004) *Measuring the Stock of Human Capital for Australia* (No. 1351.0.55.001). Australian Bureau of Statistics

Williams, R. (2006) Generalized Ordered Logit/ Partial Proportional Odds Models for Ordinal Dependent Variables. *The Stata Journal*, 6 (1): 58-82

Appendix

A1. Estimated Health Index Pooled Across all Waves - Men

Dependent variable: SAH	**Fully Constrained Model**	**Partially Constrained Model**			
		Poor vs. rest	**Poor & fair vs. rest**	**Poor to good vs. rest**	**Poor to v. good vs. rest**
GHQ	-0.062***	-0.062***	-0.064***	-0.063***	-0.056***
	-0.001	-0.002	-0.001	-0.001	-0.001
Asthma	-0.225***	-0.224***	-0.224***	-0.224***	-0.224***
	-0.019	-0.019	-0.019	-0.019	-0.019
Arthritis	-0.174***	-0.097***	-0.181***	-0.218***	-0.246***
	-0.017	-0.029	-0.022	-0.023	-0.031
Cong. heart failure	-0.212***	-0.276***	-0.314***	-0.241*	0.173
	-0.078	-0.085	-0.102	-0.136	-0.193
Cor. heart disease	-0.225***	-0.252***	-0.252***	-0.252***	-0.252***
	-0.041	-0.042	-0.042	-0.042	-0.042
Angina	-0.210***	-0.105**	-0.268***	-0.340***	-0.271***
	-0.035	-0.049	-0.046	-0.054	-0.073
Heart attack	-0.285***	-0.288***	-0.288***	-0.288***	-0.288***
	-0.027	-0.028	-0.028	-0.028	-0.028
Stroke	-0.260***	-0.173***	-0.283***	-0.390***	-0.299***
	-0.033	-0.048	-0.044	-0.051	-0.073
Emphysema	-0.469***	-0.490***	-0.490***	-0.490***	-0.490***
	-0.072	-0.073	-0.073	-0.073	-0.073
Hyperthyroidism	-0.176*	-0.177*	-0.177*	-0.177*	-0.177*
	-0.091	-0.094	-0.094	-0.094	-0.094
Hypothyroidism	-0.238***	-0.239***	-0.239***	-0.239***	-0.239***
	-0.051	-0.053	-0.053	-0.053	-0.053
Chronic bronchitis	-0.310***	-0.321***	-0.321***	-0.321***	-0.321***
	-0.052	-0.053	-0.053	-0.053	-0.053
Liver condition	-0.277***	-0.278***	-0.278***	-0.278***	-0.278***
	-0.048	-0.049	-0.049	-0.049	-0.049
Cancer	-0.399***	-0.541***	-0.409***	-0.343***	-0.313***
	-0.033	-0.049	-0.039	-0.041	-0.054
Diabetes	-0.441***	-0.303***	-0.446***	-0.570***	-0.500***
	-0.021	-0.037	-0.027	-0.028	-0.04
Epilepsy	-0.303***	-0.287***	-0.287***	-0.287***	-0.287***
	-0.062	-0.06	-0.06	-0.06	-0.06
High blood pressure	-0.243***	-0.063**	-0.197***	-0.307***	-0.369***
	-0.014	-0.028	-0.019	-0.019	-0.024
Depression	-0.208***	-0.097**	-0.188***	-0.256***	-0.282***
	-0.027	-0.042	-0.033	-0.037	-0.049
Mobility	-0.575***	-0.669***	-0.592***	-0.486***	-0.383***
	-0.018	-0.026	-0.021	-0.024	-0.039
Lifting etc.	-0.528***	-0.457***	-0.566***	-0.551***	-0.463***
	-0.017	-0.025	-0.02	-0.023	-0.037
Manual dexterity	-0.170***	-0.158***	-0.158***	-0.158***	-0.158***
	-0.026	-0.025	-0.025	-0.025	-0.025
Continence	-0.306***	-0.255***	-0.372***	-0.354***	-0.240***
	-0.03	-0.038	-0.038	-0.045	-0.073
Hearing	-0.052**	-0.064**	-0.064**	-0.064**	-0.064**
	-0.025	-0.026	-0.026	-0.026	-0.026
Sight	-0.153***	-0.055	-0.193***	-0.246***	-0.189***
	-0.029	-0.043	-0.036	-0.044	-0.065
Communication	-0.118***	-0.097**	-0.097**	-0.097**	-0.097**
	-0.046	-0.043	-0.043	-0.043	-0.043
Memory etc.	-0.236***	-0.218***	-0.218***	-0.218***	-0.218***
	-0.025	-0.024	-0.024	-0.024	-0.024
Recognising danger	-0.061	-0.039	-0.039	-0.039	-0.039
	-0.062	-0.06	-0.06	-0.06	-0.06
Balance	-0.172***	-0.167***	-0.167***	-0.167***	-0.167***

	-0.026	-0.025	-0.025	-0.025	-0.025
Personal care	-0.140***	-0.247***	-0.058	0.139***	0.245***
	-0.034	-0.036	-0.04	-0.053	-0.082
Other problem	-0.506***	-0.421***	-0.555***	-0.513***	-0.473***
	-0.017	-0.026	-0.022	-0.025	-0.04
Age	-0.033***	-0.027***	-0.027***	-0.027***	-0.027***
	-0.002	-0.002	-0.002	-0.002	-0.002
Age squared	0.000***	0.000***	0.000***	0.000***	0.000***
	0.000	0.000	0.000	0.000	0.000
Divorced/separated	-0.041**	-0.041**	-0.041**	-0.041**	-0.041**
	-0.018	-0.019	-0.019	-0.019	-0.019
Widowed	0.014	0.022	0.022	0.022	0.022
	-0.028	-0.029	-0.029	-0.029	-0.029
Single/never married	-0.068***	-0.063***	-0.063***	-0.063***	-0.063***
	-0.015	-0.015	-0.015	-0.015	-0.015
No. of kids under 16	0.006	0.008	0.008	0.008	0.008
	-0.006	-0.006	-0.006	-0.006	-0.006
Other qualification	0.120***	0.156***	0.211***	0.076***	0.029
	-0.022	-0.041	-0.028	-0.027	-0.033
GCSE	0.126***	0.170***	0.208***	0.084***	0.039
	-0.019	-0.038	-0.025	-0.023	-0.028
A-level	0.168***	0.151***	0.251***	0.144***	0.068**
	-0.019	-0.035	-0.024	-0.023	-0.027
Further Education	0.294***	0.227***	0.383***	0.292***	0.198***
	-0.023	-0.049	-0.031	-0.028	-0.032
Degree	0.387***	0.301***	0.428***	0.394***	0.304***
	-0.02	-0.042	-0.027	-0.024	-0.028
HH income	0.155***	0.107***	0.178***	0.168***	0.125***
	-0.011	-0.024	-0.015	-0.013	-0.014
Mixed ethnic group	-0.039	-0.041	-0.041	-0.041	-0.041
	-0.042	-0.041	-0.041	-0.041	-0.041
Asian/British	-0.054***	-0.118***	0.004	-0.106***	-0.038
	-0.019	-0.039	-0.026	-0.021	-0.024
Black	0.095***	0.077***	0.077***	0.077***	0.077***
	-0.027	-0.027	-0.027	-0.027	-0.027
Arabic	-0.106	-0.338***	-0.044	-0.179**	-0.066
	-0.072	-0.111	-0.104	-0.079	-0.098
Other ethnic group	-0.02	-0.019	-0.019	-0.019	-0.019
	-0.078	-0.079	-0.079	-0.079	-0.079
Self-employed	0.097***	0.090***	0.090***	0.090***	0.090***
	-0.015	-0.016	-0.016	-0.016	-0.016
Unemployed	-0.080***	-0.234***	-0.132***	-0.080***	0.042*
	-0.018	-0.036	-0.024	-0.021	-0.024
Retired	0.003	-0.016	-0.016	-0.016	-0.016
	-0.019	-0.02	-0.02	-0.02	-0.02
Other job	-0.066***	-0.473***	-0.181***	0.036*	0.138***
	-0.017	-0.03	-0.022	-0.019	-0.021
Cut- off 1	2.583***	2.731***	0.832***	-0.084	-0.762***
	-0.115	-0.239	-0.149	-0.13	-0.142
Cut- off 2	1.300***				
	-0.114				
Cut- off 3	0.15				
	-0.114				
Cut- off 4	-1.035***				
	-0.114				
N	134,675	134,675			
Log likelihood	-162,101	-160,770			
Pseudo R^2	0.172	0.179			
Wald Chi^2	22,814	25,560			
Prob > Chi^2	0.000	0.000			

Note: 1. Robust SE corrected for the clustering within-individuals in parentheses;
*2. Omitted groups: Residence: North East, North West, London, Midlands, South East, South West, Wales and Scotland (Baseline=Northern Ireland); Waves 2 to 8; * p<.1, ** p<.05, *** p<.01*

A2. Estimated Health Index Pooled Across all Waves – Women

Dependent variable: SAH	**Fully Constrained Model**	**Partially Constrained Model**			
		Poor vs. rest	**Poor & fair vs. rest**	**Poor to good vs. rest**	**Poor to v. good vs. rest**
GHQ	-0.057***	-0.055***	-0.060***	-0.057***	-0.053***
	-0.001	-0.001	-0.001	-0.001	-0.001
Asthma	-0.267***	-0.224***	-0.251***	-0.277***	-0.319***
	-0.015	-0.027	-0.02	-0.018	-0.024
Arthritis	-0.203***	-0.066***	-0.192***	-0.268***	-0.297***
	-0.013	-0.023	-0.018	-0.018	-0.024
Cong. heart failure	-0.294***	-0.344***	-0.344***	-0.344***	-0.344***
	-0.084	-0.085	-0.085	-0.085	-0.085
Cor. heart disease	-0.284***	-0.240***	-0.385***	-0.368***	-0.127
	-0.048	-0.06	-0.059	-0.081	-0.12
Angina	-0.213***	-0.274***	-0.274***	-0.274***	-0.274***
	-0.036	-0.036	-0.036	-0.036	-0.036
Heart attack	-0.278***	-0.282***	-0.282***	-0.282***	-0.282***
	-0.037	-0.037	-0.037	-0.037	-0.037
Stroke	-0.182***	-0.127***	-0.193***	-0.319***	-0.240***
	-0.035	-0.047	-0.047	-0.056	-0.077
Emphysema	-0.476***	-0.491***	-0.491***	-0.491***	-0.491***
	-0.058	-0.057	-0.057	-0.057	-0.057
Hyperthyroidism	-0.220***	-0.228***	-0.228***	-0.228***	-0.228***
	-0.04	-0.041	-0.041	-0.041	-0.041
Hypothyroidism	-0.205***	-0.082**	-0.160***	-0.250***	-0.359***
	-0.021	-0.039	-0.029	-0.029	-0.039
Chronic bronchitis	-0.303***	-0.195***	-0.505***	-0.327***	-0.390***
	-0.045	-0.061	-0.062	-0.067	-0.105
Liver condition	-0.359***	-0.240***	-0.416***	-0.476***	-0.533***
	-0.043	-0.066	-0.055	-0.061	-0.099
Cancer	-0.410***	-0.418***	-0.418***	-0.418***	-0.418***
	-0.029	-0.029	-0.029	-0.029	-0.029
Diabetes	-0.441***	-0.274***	-0.468***	-0.616***	-0.556***
	-0.021	-0.034	-0.028	-0.032	-0.051
Epilepsy	-0.313***	-0.279***	-0.279***	-0.279***	-0.279***
	-0.061	-0.059	-0.059	-0.059	-0.059
High blood pressure	-0.231***	-0.071***	-0.188***	-0.311***	-0.345***
	-0.013	-0.024	-0.018	-0.018	-0.025
Depression	-0.270***	-0.122***	-0.280***	-0.336***	-0.299***
	-0.018	-0.028	-0.023	-0.024	-0.034
Mobility	-0.625***	-0.656***	-0.668***	-0.579***	-0.412***
	-0.015	-0.022	-0.018	-0.021	-0.034
Lifting etc.	-0.534***	-0.462***	-0.552***	-0.565***	-0.560***
	-0.013	-0.021	-0.016	-0.019	-0.031
Manual dexterity	-0.204***	-0.217***	-0.217***	-0.217***	-0.217***
	-0.019	-0.019	-0.019	-0.019	-0.019
Continence	-0.205***	-0.154***	-0.272***	-0.261***	-0.275***
	-0.025	-0.032	-0.033	-0.042	-0.068
Hearing	-0.057**	-0.075***	-0.075***	-0.075***	-0.075***
	-0.026	-0.026	-0.026	-0.026	-0.026
Sight	-0.173***	-0.123***	-0.219***	-0.252***	-0.256***
	-0.025	-0.034	-0.034	-0.044	-0.068
Communication	-0.053	-0.016	-0.016	-0.016	-0.016
	-0.046	-0.044	-0.044	-0.044	-0.044
Memory etc.	-0.318***	-0.228***	-0.372***	-0.363***	-0.179***
	-0.021	-0.026	-0.026	-0.033	-0.052
Recognising danger	0.142***	0.140***	0.140***	0.140***	0.140***
	-0.052	-0.05	-0.05	-0.05	-0.05
Balance	-0.219***	-0.230***	-0.230***	-0.230***	-0.230***
	-0.021	-0.021	-0.021	-0.021	-0.021
Personal care	-0.240***	-0.354***	-0.185***	0.037	0.266***

	-0.025	-0.027	-0.032	-0.045	-0.072
Other problem	-0.572***	-0.497***	-0.600***	-0.600***	-0.515***
	-0.015	-0.021	-0.019	-0.022	-0.037
Age	-0.020***	-0.015***	-0.024***	-0.016***	-0.010***
	-0.002	-0.003	-0.002	-0.002	-0.002
Age squared	0.000***	0.000***	0.000***	0.000***	0.000**
	0.000	0.000	0.000	0.000	0.000
Divorced/separated	-0.018	-0.090***	-0.029	-0.005	0.014
	-0.014	-0.026	-0.018	-0.017	-0.02
Widowed	0.057***	0.057***	0.057***	0.057***	0.057***
	-0.019	-0.019	-0.019	-0.019	-0.019
Single/never married	-0.065***	-0.062***	-0.062***	-0.062***	-0.062***
	-0.014	-0.014	-0.014	-0.014	-0.014
No. of kids under 16	0.025***	0.044***	0.045***	0.020***	0.007
	-0.005	-0.011	-0.007	-0.006	-0.007
Other qualification	0.112***	0.119***	0.163***	0.096***	0.042
	-0.019	-0.034	-0.024	-0.024	-0.031
GCSE	0.155***	0.166***	0.215***	0.141***	0.076***
	-0.016	-0.03	-0.021	-0.02	-0.026
A-level	0.244***	0.244***	0.284***	0.253***	0.149***
	-0.017	-0.034	-0.023	-0.021	-0.026
Further Education	0.248***	0.221***	0.310***	0.254***	0.161***
	-0.018	-0.036	-0.025	-0.023	-0.028
Degree	0.374***	0.290***	0.384***	0.388***	0.299***
	-0.018	-0.036	-0.024	-0.022	-0.027
HH income	0.163***	0.062***	0.165***	0.183***	0.151***
	-0.01	-0.021	-0.013	-0.012	-0.013
Mixed ethnic group	-0.148***	-0.148***	-0.148***	-0.148***	-0.148***
	-0.034	-0.034	-0.034	-0.034	-0.034
Asian/British	-0.211***	-0.192***	-0.217***	-0.279***	-0.143***
	-0.017	-0.035	-0.023	-0.02	-0.023
Black	-0.081***	-0.127**	-0.169***	-0.125***	0.023
	-0.024	-0.054	-0.032	-0.027	-0.03
Arabic	0.028	0.012	0.012	0.012	0.012
	-0.074	-0.073	-0.073	-0.073	-0.073
Other ethnic group	-0.245***	-0.257***	-0.257***	-0.257***	-0.257***
	-0.069	-0.068	-0.068	-0.068	-0.068
Self-employed	0.097***	0.098***	0.098***	0.098***	0.098***
	-0.019	-0.019	-0.019	-0.019	-0.019
Unemployed	-0.120***	-0.254***	-0.214***	-0.110***	0.031
	-0.018	-0.038	-0.024	-0.021	-0.025
Retired	-0.019	-0.136***	-0.027	-0.002	-0.036
	-0.017	-0.034	-0.023	-0.021	-0.027
Other job	-0.102***	-0.394***	-0.209***	-0.034**	0.030*
	-0.011	-0.026	-0.016	-0.013	-0.015
Cut- off 1	2.234***	2.759***	0.930***	-0.513***	-1.416***
	-0.103	-0.213	-0.138	-0.12	-0.13
Cut- off 2	0.911***				
	-0.103				
Cut- off 3	-0.240**				
	-0.103				
Cut- off 4	-1.463***				
	-0.103				
N	170,133	170,133			
Log likelihood	-202,992	-201,162			
Pseudo R^2	0.186	0.193			
Wald Chi^2	32,461	35,898			
Prob > Chi^2	0.000	0.000			

Note: 1. Robust SE corrected for the clustering within-individuals in parentheses;
*2. Omitted groups: Residence: North East, North West, London, Midlands, South East, South West, Wales and Scotland (Baseline=Northern Ireland); Waves 2 to 8; * p<.1, ** p<.05, *** p<.01*

A3. Dynamic Retirement Model Using RE Probit Regression

Dependent variable: Retired	**Men**	**Women**
Lagged HI (t-1)	-0.016***	-0.026***
	(0.006)	(0.005)
Initial HI (t=1)	-0.028***	-0.016*
	(0.009)	(0.009)
Working spouse in good health (t-1)	-0.224*	-0.055
	(0.115)	(0.127)
Working spouse in poor health (t-1)	-0.456**	0.209
	(0.212)	(0.246)
Retired spouse in good health (t-1)	0.482***	0.871***
	(0.132)	(0.141)
Retired spouse in poor health (t-1)	0.264	0.801***
	(0.206)	(0.233)
Carer outside HH (t-1)	-0.023	0.057
	(0.090)	(0.085)
Carer within HH (t-1)	-0.004	-0.009
	(0.163)	(0.161)
Carer within and outside HH (t-1)	0.074	0.268
	(0.320)	(0.235)
Age	1.327***	1.573***
	(0.150)	(0.159)
Age squared	-0.008***	-0.010***
	(0.001)	(0.001)
Other	-0.435**	0.132
	(0.206)	(0.234)
GCSE	-0.071	0.092
	(0.202)	(0.210)
A-level	0.149	0.073
	(0.189)	(0.223)
FE	0.335	0.288
	(0.225)	(0.221)
Degree	0.489**	0.203
	(0.201)	(0.226)
Private firm/business (t=1)	0.770***	1.030***
	(0.143)	(0.212)
Government (t=1)	1.512***	1.540***
	(0.181)	(0.216)
Other job sector (t=1)	1.197***	1.246***
	(0.196)	(0.220)
Log HH income	0.221*	0.805***
	(0.117)	(0.131)
Home with mortgage/loan	-1.044***	-1.069***
	(0.123)	(0.137)
Home rented privately	-0.640***	-0.899***
	(0.183)	(0.185)
Home rented from local authority	-0.351	-1.236***
	(0.244)	(0.264)
Constant	-53.254***	-68.710***
	(5.116)	(5.459)
lnsig2u constant	1.527***	1.677***
	(0.094)	(0.094)
N	18,295	19,373
Log likelihood	-3,933.162	-3,943.745
Wald chi2	601.378	624.590
Prob > chi2	0.000	0.000

Note: 1. Robust SE corrected for the clustering within-individuals in parentheses;
2. Omitted groups: Residence: North East, North West, London, Midlands, South East, South West, Wales and Scotland (Baseline=Northern Ireland); Waves 3 to 8.
** p<.1, ** p<.05, *** p<.01*

A4. Selection Equation for the Wage Model using Probit Regression

Dependent variable: Work	**No Qual.**	**Other Qual.**	**GCSE**	**A-level**	**Further edu.**	**Degree**
Men						
Owning a house	0.094	-0.012	-0.099*	-0.248***	-0.350***	-0.372***
	(0.086)	(0.077)	(0.054)	(0.047)	(0.074)	(0.046)
Initial HI (t=1)	0.019***	0.013***	0.016***	0.020***	0.015**	0.021***
	(0.005)	(0.005)	(0.004)	(0.003)	(0.006)	(0.003)
Estimated HI	0.062***	0.054***	0.054***	0.052***	0.044***	0.036***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.005)	(0.003)
N	5,343	6,715	15,578	19,423	8,206	21,051
Log likelihood	-2,149.696	-2,736.918	-6,087.346	-7,190.994	-2,611.836	-6,245.519
Pseudi R^2	0.408	0.372	0.371	0.410	0.418	0.378
Wald chi2	804.089	872.848	2,259.852	2,657.140	818.076	2,071.969
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000

Women						
No. of kids	-0.637***	-0.584***	-0.642***	-0.567***	-0.695***	-0.716***
	(0.119)	(0.095)	(0.038)	(0.037)	(0.051)	(0.034)
Initial HI (t=1)	0.021***	0.022***	0.023***	0.015***	0.021***	0.018***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Estimated HI	0.050***	0.045***	0.046***	0.044***	0.037***	0.040***
	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)
N	8,440	7,608	21,987	20,531	13,388	26,162
Log likelihood	-3,333.841	-3,560.342	-10,731.997	-9,933.123	-5,899.269	-10,978.061
Pseudi R^2	0.298	0.324	0.286	0.276	0.293	0.222
Wald chi2	633.771	842.321	2,345.936	2,430.415	1,423.298	1,979.868
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000

Note: *1. Robust SE corrected for clustering within-individuals in parentheses;*
2. full regression results available from authors upon request.
** p<.1, ** p<.05, *** p<.01*

A5. Wage Equation Using OLS

	Men						Women					
Dependent variable: Log hourly wage	**No Qual.**	**Other Qual.**	**GCSE**	**A-level**	**Further edu.**	**Degree**	**No Qual.**	**Other Qual.**	**GCSE**	**A-level**	**Further edu.**	**Degree**
IRM	0.186	0.184	0.257**	0.063	-0.116	0.008	0.070	0.308**	0.143**	0.117*	0.086	0.128**
	(0.157)	(0.159)	(0.104)	(0.080)	(0.107)	(0.101)	(0.119)	(0.128)	(0.063)	(0.068)	(0.075)	(0.051)
Lagged HI (t-1)	0.007	0.010**	0.011***	0.008***	0.004	0.007***	0.001	0.010**	0.005***	0.007***	0.004**	0.007***
	(0.006)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)	(0.005)	(0.004)	(0.002)	(0.002)	(0.002)	(0.001)
Initial HI (t=1)	0.006*	0.002	0.008***	0.006***	0.002	0.007***	0.005	0.008**	0.008***	-0.001	0.001	0.003**
	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Working spouse in good health (t-1)	0.150**	0.057	0.031	0.034*	0.004	0.080***	-0.015	0.040	-0.026	-0.026	0.024	0.001
	(0.065)	(0.042)	(0.021)	(0.020)	(0.028)	(0.019)	(0.033)	(0.029)	(0.017)	(0.017)	(0.020)	(0.017)
Working spouse in poor health (t-1)	0.165*	-0.014	0.004	-0.033	-0.075	-0.029	-0.107**	0.020	-0.104***	-0.027	0.054	-0.050
	(0.086)	(0.051)	(0.055)	(0.040)	(0.085)	(0.070)	(0.049)	(0.053)	(0.040)	(0.050)	(0.067)	(0.044)
Retired spouse in good health (t-1)	0.082	-0.072	-0.067	0.006	-0.034	0.347***	-0.046	-0.031	-0.058	-0.093	-0.002	-0.002
	(0.081)	(0.087)	(0.114)	(0.110)	(0.086)	(0.110)	(0.082)	(0.100)	(0.054)	(0.070)	(0.072)	(0.061)
Retired spouse in poor health (t-1)	-0.059	-0.177	-0.010	0.094	-0.388	0.114	-0.024	-0.170	0.051	-0.217**	0.164	0.033
	(0.167)	(0.111)	(0.101)	(0.143)	(0.237)	(0.104)	(0.108)	(0.154)	(0.111)	(0.096)	(0.138)	(0.082)
Carer within HH (t-1)	-0.128	0.009	-0.071	-0.141***	-0.043	-0.017	-0.014	-0.196***	-0.098***	-0.113***	-0.110**	-0.020
	(0.078)	(0.057)	(0.044)	(0.036)	(0.073)	(0.042)	(0.086)	(0.063)	(0.033)	(0.032)	(0.052)	(0.031)
Carer outside HH (t-1)	0.092*	0.035	-0.016	-0.053**	0.013	-0.052**	0.062	0.056*	-0.031*	-0.012	0.027	-0.009
	(0.051)	(0.034)	(0.025)	(0.025)	(0.032)	(0.023)	(0.040)	(0.032)	(0.017)	(0.016)	(0.020)	(0.017)
Carer within and outside HH (t-1)	-0.351**	-0.006	-0.132	-0.026	0.174	0.099	0.072	0.156	-0.050	-0.155**	-0.091	-0.172**
	(0.156)	(0.100)	(0.091)	(0.071)	(0.156)	(0.101)	(0.062)	(0.126)	(0.088)	(0.061)	(0.066)	(0.067)
Age	0.030	0.078***	0.072***	0.063***	0.044***	0.061***	0.017	0.077***	0.037***	0.048***	0.040***	0.062***
	(0.025)	(0.020)	(0.014)	(0.012)	(0.015)	(0.012)	(0.024)	(0.026)	(0.009)	(0.009)	(0.010)	(0.007)
Age squared	-0.000	-0.001***	-0.001***	-0.001***	-0.000**	-0.001***	-0.000	-0.001***	-0.000***	-0.001***	-0.000***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Private firm/business (t=1)	0.168	0.004	0.131*	0.112**	0.172**	0.099*	0.198**	-0.052	0.050	0.156**	0.046	0.134*
	(0.131)	(0.128)	(0.077)	(0.045)	(0.074)	(0.056)	(0.093)	(0.159)	(0.066)	(0.070)	(0.094)	(0.074)
Government (t=1)	0.218	-0.019	0.134*	0.140***	0.134*	0.030	0.318***	0.047	0.107	0.263***	0.150	0.185**
	(0.147)	(0.129)	(0.078)	(0.048)	(0.077)	(0.057)	(0.104)	(0.164)	(0.067)	(0.070)	(0.095)	(0.073)
Other job sector (t=1)	0.138	-0.015	0.215**	0.132**	0.115	0.057	0.267***	0.106	0.100	0.236***	0.213**	0.193***
	(0.148)	(0.135)	(0.086)	(0.054)	(0.080)	(0.058)	(0.100)	(0.162)	(0.068)	(0.071)	(0.095)	(0.074)
PT	-0.064	-0.051	-0.121***	-0.178***	-0.009	0.078	-0.005	-0.050	-0.059***	-0.041***	-0.026	-0.004
	(0.060)	(0.054)	(0.038)	(0.043)	(0.063)	(0.048)	(0.032)	(0.031)	(0.014)	(0.015)	(0.020)	(0.018)
Experience	0.016***	0.013***	0.007*	0.006	-0.002	0.014***	0.004	0.005	0.014***	0.008*	0.007	0.010***
	(0.006)	(0.004)	(0.004)	(0.005)	(0.006)	(0.004)	(0.006)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)
Experience squared	-0.035**	-0.016*	-0.004	-0.002	0.010	-0.032**	-0.001	-0.014	-0.021**	-0.005	-0.000	-0.029**

	(0.015)	(0.010)	(0.012)	(0.013)	(0.017)	(0.015)	(0.018)	(0.014)	(0.009)	(0.014)	(0.011)	(0.012)
Manager	0.182	0.361***	0.372***	0.685***	0.472***	0.612***	0.373***	0.449***	0.380***	0.378***	0.484***	0.560***
	(0.152)	(0.105)	(0.078)	(0.137)	(0.158)	(0.136)	(0.131)	(0.085)	(0.082)	(0.103)	(0.054)	(0.043)
Professional	0.323*	0.508***	0.554***	0.630***	0.417***	0.618***	0.217*	0.726***	0.535***	0.460***	0.394***	0.639***
	(0.173)	(0.154)	(0.104)	(0.144)	(0.160)	(0.138)	(0.122)	(0.211)	(0.152)	(0.122)	(0.083)	(0.045)
Non-manual	-0.113	0.224**	0.238***	0.569***	0.307**	0.437***	0.165*	0.151***	0.169**	0.167*	0.319***	0.357***
	(0.128)	(0.096)	(0.075)	(0.136)	(0.156)	(0.137)	(0.088)	(0.030)	(0.076)	(0.101)	(0.051)	(0.040)
Skilled-manual	-0.121	0.131	0.182**	0.463***	0.151	0.177	0.027	-0.002	-0.014	0.014	0.062	0.023
	(0.130)	(0.093)	(0.073)	(0.135)	(0.160)	(0.141)	(0.089)	(0.034)	(0.077)	(0.102)	(0.054)	(0.048)
Unskilled	-0.207	0.031	0.028	0.297**	0.000	0.016	0.105		-0.005	-0.024	0.109**	
	(0.128)	(0.095)	(0.074)	(0.136)	(0.161)	(0.143)	(0.085)		(0.076)	(0.102)	(0.055)	
3-9 employees	0.010	0.040	0.072	0.138*	0.172*	0.127	(0.219)	(0.126)	(0.085)	(0.086)	(0.134)	(0.108)
	(0.106)	(0.092)	(0.065)	(0.079)	(0.090)	(0.078)	-0.014	-0.155*	0.059	-0.075	-0.082	0.033
10-24 employees	0.019	0.190**	0.163**	0.154*	0.284***	0.215***	(0.070)	(0.092)	(0.058)	(0.058)	(0.079)	(0.074)
	(0.105)	(0.093)	(0.065)	(0.081)	(0.088)	(0.074)	0.025	-0.193**	0.096*	-0.040	0.015	0.114
25-49 employees	0.158	0.215**	0.142**	0.180**	0.303***	0.284***	(0.066)	(0.089)	(0.056)	(0.058)	(0.077)	(0.074)
	(0.119)	(0.098)	(0.064)	(0.081)	(0.089)	(0.073)	0.079	-0.076	0.032	-0.046	0.051	0.174**
50-99 employees	0.161	0.170*	0.197***	0.253***	0.434***	0.305***	(0.069)	(0.102)	(0.057)	(0.058)	(0.077)	(0.074)
	(0.111)	(0.097)	(0.066)	(0.082)	(0.088)	(0.074)	0.171*	-0.172*	0.078	-0.005	0.073	0.220***
100-199 employees	0.249**	0.267***	0.209***	0.219***	0.470***	0.313***	(0.090)	(0.090)	(0.059)	(0.060)	(0.079)	(0.074)
	(0.112)	(0.095)	(0.066)	(0.081)	(0.091)	(0.073)	0.159**	-0.128	0.101*	0.078	0.079	0.208***
200-499 employees	0.134	0.305***	0.279***	0.281***	0.434***	0.363***	(0.078)	(0.096)	(0.060)	(0.060)	(0.081)	(0.073)
	(0.107)	(0.096)	(0.066)	(0.081)	(0.090)	(0.073)	0.052	-0.135	0.099*	0.061	0.139*	0.241***
500-999 employees	0.221*	0.252***	0.340***	0.297***	0.476***	0.423***	(0.075)	(0.094)	(0.059)	(0.060)	(0.079)	(0.075)
	(0.118)	(0.096)	(0.068)	(0.085)	(0.095)	(0.074)	0.110	-0.082	0.178***	0.123*	0.158*	0.281***
1000 or more employees	0.207*	0.373***	0.412***	0.416***	0.552***	0.482***	(0.094)	(0.105)	(0.063)	(0.065)	(0.081)	(0.075)
	(0.113)	(0.100)	(0.068)	(0.084)	(0.088)	(0.072)	0.225***	-0.082	0.265***	0.179***	0.216***	0.340***
Unknown but <25	-0.128	0.000	0.295***	0.885**	-0.001	0.246***	0.196**	0.053	0.097	-0.176*	0.253*	0.428**
	(0.134)	(.)	(0.113)	(0.361)	(0.194)	(0.084)	(0.097)	(0.115)	(0.175)	(0.099)	(0.143)	(0.197)
Unknown but >=25	0.646***	0.174	0.314***	0.215*	0.182	0.239**	0.266	-0.329***	0.137	-0.030	0.335**	0.144
	(0.207)	(0.119)	(0.096)	(0.119)	(0.226)	(0.121)	(0.219)	(0.126)	(0.085)	(0.086)	(0.134)	(0.108)
Constant	-0.056	-1.093	-1.542**	-1.216**	0.057	-1.152**	0.582	-1.385	-0.231	0.289	0.548	-0.568**
	(1.148)	(0.904)	(0.617)	(0.495)	(0.585)	(0.450)	(1.100)	(1.088)	(0.431)	(0.412)	(0.431)	(0.284)
N	880	1,982	4,877	5,749	3,196	8,930	1,144	1,859	6,547	6,170	5,320	10,646
Adjusted R^2	0.248	0.225	0.279	0.301	0.270	0.287	0.138	0.244	0.247	0.271	0.257	0.269

Note: 1. Robust SE corrected for the clustering within-individuals in parentheses;

2. Omitted groups: Residence: North East, North West, London, Midlands, South East, South West, Wales and Scotland (Baseline=Northern Ireland); Waves 3 to 8.

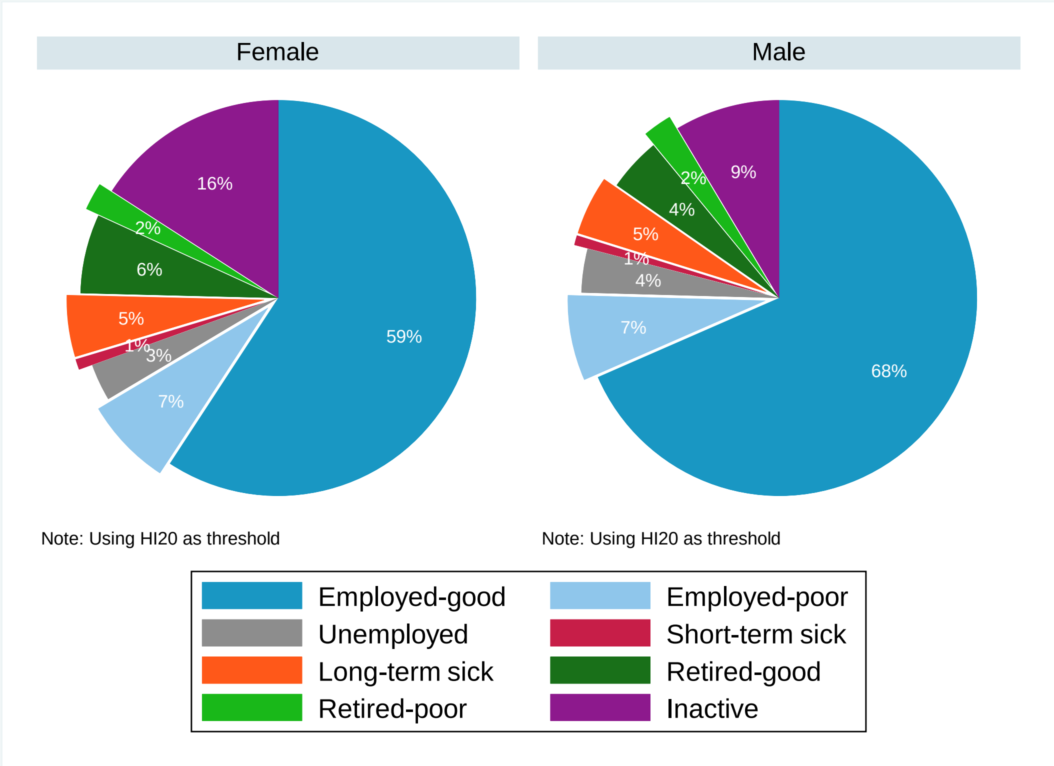
** p<.1, ** p<.05, *** p<.01*

A6. Summary of Variables used in the HCS Estimation

Variable	Classification	Years	Source
Annual earnings	Gender, age, qualification and health	1997-18, 2010-18	Quarterly LFS, UKHLS
Survival rate	Gender and age	1997-17	ONS
Employment rate	Gender, age and qualification	1997-18	Quarterly LFS
Unemployment rate	Gender, age and qualification	1997-18	Quarterly LFS
Retirement rate	Gender, age, qualification and health	1997-18, 2010-18	Quarterly LFS, UKHLS
Short-term sickness rate	Gender, age and qualification	1997-18	Quarterly LFS
Long-term sickness rate	Gender, age and qualification	1997-18	Quarterly LFS
Other inactivity rate	Gender, age and qualification	1997-18	Quarterly LFS
Enrolment rate in school	Gender, age and qualification	1997-17	Longitudinal LFS
Enrolment rate in further education	Gender, age and qualification	1997-17	Longitudinal LFS
Enrolment rate in higher education	Gender, age and qualification	1997-17	Longitudinal LFS
Population	Gender, age, qualification and health	1997-18, 2010-18	Quarterly LFS, ONS, UKHLS

A7. Population Aged 16-69 by Job and Health Status (in per cent), using HI20 as threshold

A7a. 2018

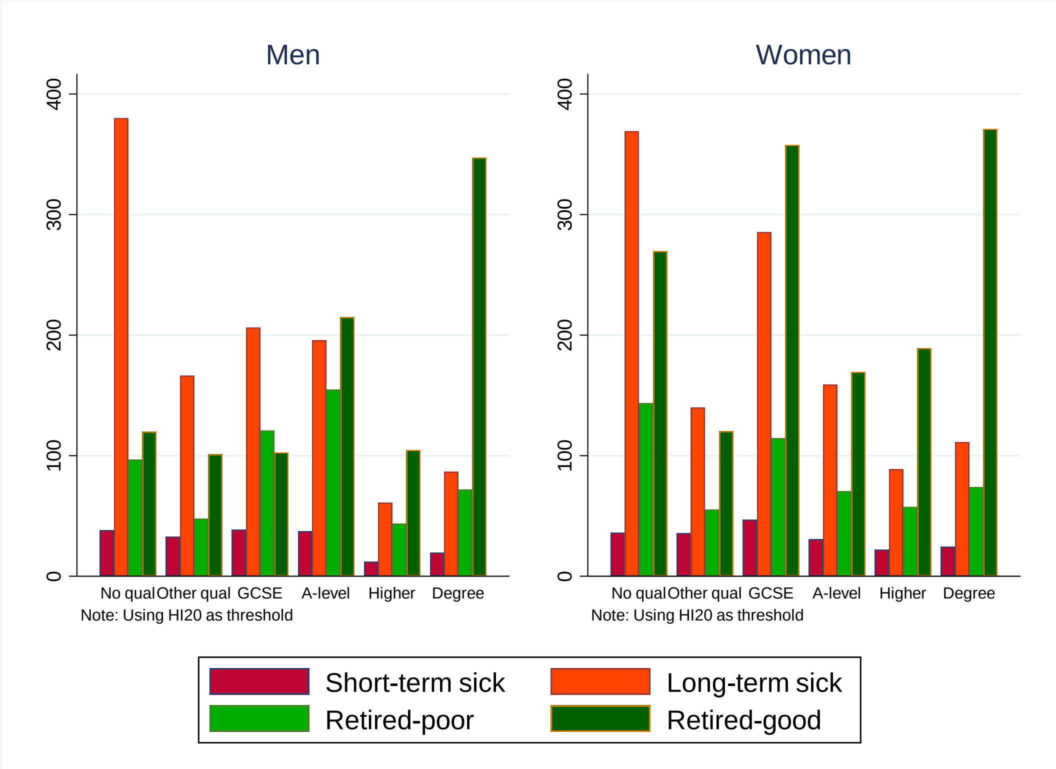


A7b. 1996-2018

Note: Using HI20 as threshold

Source: UKHLS, LFS and authors' own calculations

A8. Population Aged 16-69 by Qualification and Health Status (in thousand), 2018, using HI20 as threshold



Source: UKHLS, LFS and authors' own calculations

A9. Sickness, Presenteeism and Absenteeism, 1996-2018, using HI20 as threshold

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Total																							
Productive HCS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Potential HCS	**1.65**	**1.62**	**1.60**	**1.60**	**1.57**	**1.57**	**1.59**	**1.57**	**1.56**	**1.55**	**1.56**	**1.56**	**1.55**	**1.61**	**1.61**	**1.63**	**1.63**	**1.61**	**1.58**	**1.56**	**1.56**	**1.55**	**1.54**
Difference due to (%)																							
Long-term ill	17.7	18.7	19.7	19.0	19.5	19.3	18.9	18.6	19.2	17.9	16.9	17.4	17.4	15.8	15.6	14.9	14.4	14.6	15.1	15.5	15.4	15.6	16.2
Short-term ill	2.9	2.9	2.8	2.7	3.0	2.7	2.7	2.9	2.7	2.7	2.9	3.0	2.9	2.7	2.7	2.7	2.8	3.0	3.0	3.1	3.1	3.2	3.2
Earnings (ill-health)	0.7	1.5	2.0	2.2	2.4	2.5	2.5	2.5	2.5	2.7	2.6	2.5	2.5	2.4	2.7	2.6	2.5	2.3	2.4	2.5	2.7	2.8	2.9
Retirement (ill-health)	0.5	1.2	1.5	1.8	1.9	2.2	2.2	2.2	2.2	2.3	2.3	2.1	2.0	2.0	2.2	2.1	2.2	2.1	2.1	2.2	2.3	2.4	2.5
Retirement (general)	20.5	22.8	23.7	24.9	25.7	26.9	27.3	27.2	27.6	27.8	28.3	26.8	26.5	25.2	24.6	23.8	25.2	26.0	27.2	28.1	28.1	29.7	29.5
Other	57.7	52.9	50.4	49.4	47.3	46.5	46.4	46.6	45.9	46.7	47.0	48.2	48.7	52.0	52.2	53.9	52.9	52.1	50.2	48.6	48.4	46.3	45.7
Men																							
Productive HCS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Potential HCS	**1.51**	**1.49**	**1.46**	**1.47**	**1.44**	**1.44**	**1.46**	**1.44**	**1.42**	**1.42**	**1.44**	**1.43**	**1.42**	**1.49**	**1.49**	**1.50**	**1.50**	**1.50**	**1.47**	**1.45**	**1.44**	**1.44**	**1.44**
Difference due to (%)																							
Long-term ill	23.4	24.7	26.5	25.7	26.5	25.6	25.2	25.1	25.8	24.4	22.6	23.0	23.1	20.4	20.1	19.4	18.3	18.4	18.9	18.8	19.2	19.5	19.7
Short-term ill	3.5	3.4	3.3	3.1	3.5	3.2	3.3	3.6	3.5	3.4	3.7	3.7	3.5	3.1	3.2	3.1	3.4	3.5	3.5	3.7	3.8	3.8	3.8
Earnings (ill-health)	1.0	2.2	2.9	3.3	3.7	3.7	3.7	3.8	3.9	4.1	3.8	3.8	3.8	3.4	4.0	3.7	3.6	3.4	3.5	3.7	4.0	4.1	4.1
Retirement (ill-health)	0.5	1.2	1.5	1.8	2.0	2.2	2.3	2.2	2.2	2.3	2.4	2.3	2.1	2.0	2.3	2.2	2.3	2.2	2.3	2.4	2.5	2.6	2.7
Retirement (general)	20.1	24.0	25.1	25.9	27.2	28.7	29.3	27.8	28.6	28.9	31.0	29.7	28.1	25.9	26.0	25.6	27.4	28.9	30.7	32.4	31.8	32.9	33.4
Other	51.5	44.4	40.7	40.2	37.1	36.7	36.2	37.5	36.0	37.0	36.5	37.5	39.5	45.2	44.4	45.9	44.9	43.5	41.1	39.1	38.6	37.1	36.4
Women																							
Productive HCS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Potential HCS	**1.97**	**1.94**	**1.93**	**1.92**	**1.86**	**1.90**	**1.87**	**1.86**	**1.83**	**1.79**	**1.82**	**1.80**	**1.81**	**1.84**	**1.85**	**1.89**	**1.87**	**1.82**	**1.79**	**1.77**	**1.78**	**1.78**	**1.74**
Difference due to (%)																							
Long-term ill	10.5	11.1	11.2	11.1	11.2	11.8	11.3	11.3	12.0	11.1	10.6	11.6	11.4	10.6	10.5	9.8	10.1	10.2	10.8	11.7	11.3	11.5	12.4
Short-term ill	2.3	2.3	2.1	2.1	2.4	2.0	2.1	2.1	1.9	1.9	2.0	2.2	2.4	2.2	2.1	2.2	2.1	2.3	2.3	2.5	2.4	2.5	2.5
Earnings (ill-health)	0.2	0.6	0.8	0.9	0.9	1.1	1.0	1.0	1.0	1.1	1.1	1.1	1.1	1.1	1.2	1.2	1.1	1.1	1.1	1.2	1.3	1.4	1.6
Retirement (ill-health)	0.5	1.2	1.6	1.8	1.9	2.3	2.1	2.2	2.2	2.3	2.2	2.0	2.0	2.1	2.2	2.1	2.0	1.9	1.9	1.9	2.0	2.3	2.3
Retirement (general)	20.8	21.4	22.0	23.8	24.0	24.8	24.9	26.5	26.5	26.7	25.4	23.8	24.8	24.3	23.0	21.8	22.8	22.5	23.3	23.3	24.0	26.2	25.2
Other	65.7	63.5	62.3	60.3	59.5	58.0	58.6	56.9	56.5	56.8	58.7	59.3	58.5	59.7	61.1	62.9	61.7	62.1	60.7	59.5	58.9	56.0	56.0

Source: UKHLS, LFS and authors' own calculations

A10. Sickness, Presenteeism and Absenteeism by Gender, Qualification and Age Group, 2018, using HI20 as threshold

In 2018	Low qualification			High qualification		
	F	M	T	F	M	T
			Aged 30+			
Potential HCS	**2.34**	**1.59**	**1.81**	**1.60**	**1.35**	**1.45**
Long-term ill	1.21	1.16	1.17	1.05	1.04	1.04
Short-term ill	1.03	1.03	1.03	1.01	1.01	1.01
Earnings (ill-health)	1.04	1.04	1.04	1.01	1.02	1.01
Retirement (ill-health)	1.07	1.03	1.04	1.02	1.01	1.01
Retirement (general)	1.31	1.18	1.22	1.20	1.17	1.19
			Aged 50+			
Potential HCS	**2.88**	**2.14**	**2.40**	**2.24**	**1.90**	**2.02**
Long-term ill	1.31	1.28	1.29	1.11	1.07	1.09
Short-term ill	1.04	1.03	1.04	1.02	1.01	1.01
Earnings (ill-health)	1.07	1.06	1.07	1.02	1.03	1.02
Retirement (ill-health)	1.19	1.14	1.16	1.09	1.07	1.07
Retirement (general)	1.76	1.45	1.56	1.66	1.54	1.58

Source: UKHLS, LFS and authors' own calculations