Education and its Effects on Income and Mortality of Men aged Sixty-five and over in Great Britain

Dorsett, R., Lui, S. and Weale, M.

Education and its Effects on Income and Mortality of Men aged Sixty-five and over in Great Britain

Silvia Lui and Martin Weale*†
National Institute of Economic and Social Research,
2, Dean Trench Street,
London SW1P 3HE.
United Kingdom

Revised 2nd February 2014

Abstract

We explore the effects of income and, additionally education on the income, self-reported health and survival of people aged sixty-five and over in Great Britain in order to identify benefits resulting from education which are omitted in the conventional analysis with its focus on labour income excluding employer contributions. We find, for men, that income at the age of sixty-five is significantly influenced by educational attainment and has a significant effect on survival. Even after controlling for circumstances at age sixty-five or when first observed, we identify benefits discounted to age sixty-five of £115,000 for men with higher education qualifications as compared to those with minimal qualifications.

JEL Codes: C33, C35, J17, J24

Keywords: Returns to Education; Income, Education and Life Expectancy; Mortality

7958 Words

*Corresponding author. Email mweale@niesr.ac.uk. Tel +44 207 654 1945
†We are grateful to the Economic and Social Research Council for financial support.

1 Introduction

The purpose of this paper is to explore a hitherto neglected component of the return to education for men- the benefits which may result from it post-retirement. There are two aspects to this. One is that there is a well-established link between mortality and education- although a meta-analysis (Baker et al. 2011) suggests that the effects are appreciably less clear for old people than for younger people. Secondly, there may be elements of income post-retirement which are omitted from the conventional analysis of the relationship between wages and education.

The connections between education, income and mortality are generally regarded as well-established (Smith 1999, Marmot Review 2010), having been explored in a number of different ways. Economou & Theodossiou (2011) find that, for people aged forty-five to sixty-five, both education and income affect health status, after using instrumental variables to correct for the possible role of health as a driver of income. Silles (2009) finds a clear causative influence of education on health for people aged sixty or under. Other studies look at the effects of lottery winnings (Lindahl 2005) and German unification (Frijters et al. 2005). Lleras-Muney (2005) suggests that one year of education raises life expectancy at the age of thirty-five by up to 1.7 years, based on an assessment of the effects of different compulsory education laws in different states in the United States.

Doubt is, however, cast on these findings by Clark & Royer (2013). They study the impact of increased years of education as a result of the changes to the school-leaving age in 1947 and 1972 in Great Britain. Comparing the mortality patterns of the cohorts affected by the change with those too old to be affected, they find that, if anything, the increase in compulsory education in 1947 was associated with slightly increased mortality rates between the ages of forty-five and sixty-nine.

Barker et al. (2002) argue that adult disease is strongly influenced by foetal experience, although in studies of twins both Fujiwara & Kawachi (2009) and Madsen et al. (2009) find that education plays a separate role as a determinant of adult health. Gould et al. (2011) show the importance of childhood circumstances on adult outcomes. Case & Paxson (2011) establish a link between birth-weight, childhood health subsequent career success. Related work shows a connection between childhood factors and subsequent mortality. Thus Whalley & Deary (2001) find a link between IQ at age eleven and the risk of death before the age of seventy-six but, in the absence of other control variables, this of course does not say anything about the possible magnitude of income and education effects. Batty et al. (2006) find that the effects of income on mortality are attenuated but not removed if one takes account of respondents' IQ measured at the age of fifty-six. But of course this, itself, may be a consequence of past education and income. Lager et al. (2009) find, on taking account of childhood IQ, education and income that the latter two that the ability of the latter to explain health and mortality is not much affected by the inclusion of childhood IQ as an explanatory variable. Eide & Showalter (2011) survey the field, suggesting that results typically depend on the way in which possible individual effects are treated.

Separate from, but closely related to possible relationships between education, income and mortality, the GB Census and associated longitudinal study makes it possible to compute life expectancy as a function of social class. The Census records people's occupations from which social class is derived; the longitudinal study is a one per cent sample of the Census which is linked to death records, making it possible to calculate the risk of death by age as a function of social class. The results are presented as five-year averages and show that, at the age of sixty-five a professional man had a life expectancy five years longer than that of a manual worker in the period 1997-2001. Unless education has no influence on social class, this suggests strongly that education affects life expectancy in old age.

In this paper we use the British Household Panel Survey to explore the relationship between educational status, income and survival in Great Britain for men aged sixty-five and over. For this age group income is unlikely to be strongly influenced by current health status, although it may of course be influenced by past health status which may bear on current health status. Income is likely to be strongly influenced by past education, because people are likely to receive pensions which reflect their past earnings. While the English Longitudinal Survey of Ageing also provides information on the variables of interest to us, and covers a larger initial sample, it is carried out only every two years, so has now run for five waves as compared to the sixteen available to us from the British Household Panel Survey.

We estimate a system of equations which jointly explains continuing response to the British Household Panel Survey, income at the age of sixty-five (or when first observed by the British Household Panel Survey) and survival. About half of the population of interest to us reported no significant qualifications; for the remainder we observe a relationship between education and income at sixty-five. Smoking behaviour, region of residence and health status at the age of sixty-five or when first reported, are used as control variables, as is whether there is a working member of the household. We compare our results with those of a reduced form model which examines the direct influence of educational attainment on survival.

After estimating our system of equations , we then simulate it to establish the effects of education on life expectancy at the age of sixty-five. While we do not study the dynamics of income in our panel, the relationship between age on joining the panel and income is used to infer the way in which income changes with age. This allows us also to produce an estimate of income discounted to age sixty-five. Applying a standard estimate of the value of a life year to discounted life expectancy and combining it appropriately with discounted income, it then becomes possible to value the benefit of education, in terms of longer life and higher income in old age.

2 The Data

The British Household Panel Survey (BHPS) started in 1991[1]. It is an annual survey that provides a panel of socio-economic data set over time. It interviewed each member of a household aged sixteen and over, from an initial sample of over five thousand households. The same household members are then re-interviewed in the following waves. If a member leaves the original sample household, that person, as well as the other members of the new household (aged 16 and over) are recruited for the panel. New households are also included in the survey each year in order to compensate for attrition. Deaths and non-responses are recorded. Our interest centered on the following information the BHPS provides.

1. Data on both parents' occupations when the respondents were fourteen. These identify nine one digit SOC groups. We take the 1990 SOC groups and aggregate them further into three categories:

   1) managers and administrators, professional occupations and associate professional and technical occupations

   2) clerical and secretarial occupations, craft and related occupations, personal and protective service occupations and sales occupations

   3) plant and machinery operatives, other occupations including armed forces

[1] University of Essex. Institute for Social and Economic Research, British Household Panel Survey: Waves 1-17, 1991-2008 [computer file]. 6th Edition. Colchester, Essex: UK Data Archive [distributor], May 2009. SN: 5151. It does not cover households in Northern Ireland.

Data are not available for a substantial proportion of respondents; rather than discard these observations, we use non-reporting as our baseline when looking at the influence of father's social class.

2. The response to the question on self-assessed health, "Please think back over the last 12 months about how your health has been. Compared to people of your own age, would you say that your health has on the whole been...?" Respondents are requested to report "Excellent", "Good", "Fair", "Poor", or "Very Poor". Although this is a question about relative health, the results presented by Khoman et al. (2008) suggested it could be interpreted as a proxy for a question on absolute health. In order to avoid the numerical problems which would arise if we attempted to estimate an ordered probit model to explain health status as part of our system, we aggregate the health categories, treating someone who reports their health as fair, good or excellent as having good health, with the remaining population regarded as having poor health.

3. Whether an individual did not respond or was reported dead.

4. Information on household income; this is described in more detail below.

5. Whether someone in the household is working or not.

6. The response of an individual to the question "Do you smoke cigarettes?" Respondents are required to report "Yes" or "No".

7. Information on qualifications; this is also set out in more detail below.

We were interested in the penultimate question because smoking is generally believed to be an important determinant of mortality; it was nevertheless not included in the variables considered by Contoyannis et al. (2004) in their study.

Non-response and death are recorded in BHPS, in the variable that states "Individual interview outcome"[2] that is recorded in both the data set that contains individual-level data for respondents (i.e. record type wINDRESP) and the data set that contains individual-level data for issued households (i.e. record type wINDSAMP). The former, although containing individuals' responses to the questions of our interest, covers only individuals who were actually interviewed (either in full, by proxy or by telephone). In

[2] This is given by variable wIVFIO.

order to obtain full information on respondents, non-respondents, and those reported dead, we merged the two data sets.

We look at household income at age sixty-five or when first observed if later. Equivalent household income, rather than individual income, is the appropriate variable since this influences living standards and may therefore bear on survival. The BHPS provides a gross measure of the household income. However, the net measure of household income is more appropriate for our purposes (see Jenkins (2010)). We therefore use the unofficial supplement to the income variables in the official BHPS release, the "British Household Panel Survey Derived Current and Annual Net Household Income Variables, BHPS waves 1-16, 1991-2007" constructed by the Institute of Social and Economic Research, University of Essex (see Levy & Jenkins (2008)) in our analysis. This supplementary data set contains information for those BHPS households in which all eligible household members have participated in a full interview. Those households in which one or more members refused to participate in the BHPS or whose information were given by a proxy respondents are excluded. The data set provides estimated currently weekly household net income and annual household net income for each wave. It also provides variables that classify individuals according to their family type and economic status of their family. For more detail, see Levy & Jenkins (2008). Current weekly household net income and annual household net income are recorded in the variable "whhnetde2" and "whhnyrde2" in the ISER supplement, respectively. Both variables measure total household net income which is equivalised using the Modified OECD scale (with a single adult counting as one person and a couple as 1.5 people) to adjust for differences in household composition and size. The variables are also adjusted to January 2008 prices using the Retail Price Index.

The data on educational attainment in the survey are very detailed. These were classified to match the national scale which ranges from 0 (for those with no or only minimal qualifications) to 5 for those with post-graduate degrees. The system was originally designed to represent national vocational qualifications (NVQs) but academic qualifications have also been calibrated against it, allowing most qualifications to be represented on an equal basis. We have aggregated post-graduate qualifications with other forms of higher education to give the classification of qualifications shown in table 1.

To construct our sample, we merge, wave by wave, the combined wINDRESP and wINDSAMP data set of the BHPS from above to the ISER supplement using the house-

Level 1
Youth training certificate
Trade apprenticeship
Clerical and commercial qualifications
City and Guilds Certification Part I
NVQ/SVQ level 1/SCOTVEC National Certificate Modules
GCSEs
SCEs grade D-E or 4-5
O grades A-C or 1-3
Standard grades 4-7
CSEs
O-levels (pre-1975), OLs (post-1975)
SLCs
Level 2
City and Guilds Certification Part II
NVQ/SVQ level 2/SCOTVEC Higher National Units
CPVE
1 A level
Standard grades 1-3
GNVQ
AS level
School Certificate or Matriculation
1 Higher School Certificate
Level 3
City and Guilds Certification Part III
ONC, OND, BEC/TEC/BTEC General Certificate
NVQ/SVQ level 3/SCOTVEC National Certificate or Diploma
2 or more A levels
2 or more Higher School Certificates
Higher grades
Certificate of 6th year studies
Level 4
HNC, HND, BEC/TEC/BTEC
NVQ/SVQ level 4 or 5/SCOTVEC Higher Certificate or Higher Diploma
Nursing qualifications (e.g. SEN, SRN, SCM, RGN)
Teaching qualification
University diploma or Foundation degree
University or CNAA First Degree (e.g. BA, B.Ed, BSc)
University or CNAA Higher Degree (e.g. MSc, PhD)
Other
Other post-school qualification
Total

Table 1: The Classification of Qualifications

hold identifier. Since the last available wave we consider in the ISER supplement is wave 16 (year 2006), our study thus uses the data of original sample members (OSM) between 1991 to 2006, for the period when they were aged sixty-five or older.

3 Mortality, Income and Education in the British Household Panel Survey

Our data set describes 1,260 men. Including 328 cases of recorded non-response there are a total of 8,509 observations of these men. 681 of the men are aged sixty-five or over in 1992 and thus are recorded from the beginning. The remaining 579 join during the course of the survey, with accruals varying between a low of thirty-two in 2002 and a high of fifty-eight in 1993. 575 of the men in the survey were sixty-five in the year before they were included in the data set. The average age at which men join our data set is 69.8. While it is more convenient, in our subsequent analysis, to work in terms of survival and response, we present here the data in terms of mortality and non-response because the patterns are clearer.

3.1 Influences on Mortality Rates

Here we present our data in a way which illustrates the factors influencing our subsequent modelling of income and mortality. Table 2 shows mortality rates by age for those whose incomes at sixty-five or when first observed if older, were at or below the median for the year in question, as compared to those whose incomes were above the median. At all ages those with incomes above the median had mortality rates lower than those with incomes at or below the median. To the extent that education influences income at the age of sixty-five, and to the extent that this observation is robust to questions of simultaneity of income and mortality, this table summarises the relationship explored in this paper.

Although the income/mortality relationship is at the core of this paper, it is necessary to control for other influences on mortality. Smoking is an important influence as table 2 shows. Mortality rates are higher for men who smoke than for those who do not smoke. For those aged eighty-five and over the gap is small and probably not statistically significant; relatively few smokers survive to this age.

Self-reported health status is generally thought to be another good predictor of mortality; that is borne out by table 3. This shows that, the better is self-reported health

Age		Initial Income		Smoker		Freq.
		<=Med	>Med	No	Yes	
65-69	Mean	0.025	0.014	0.014	0.033	
	S.E.	(0.005)	(0.003)	(0.003)	(0.008)	
	N	951	1,475	1,947	479	2,426
70-74	Mean	0.042	0.023	0.027	0.056	
	S.E.	(0.006)	(0.004)	(0.004)	(0.012)	
	N	1,067	1,181	1,871	377	2,248
75-79	Mean	0.054	0.042	0.040	0.101	
	S.E.	(0.007)	(0.007)	(0.005)	(0.020)	
	N	926	762	1,451	237	1,688
80-84	Mean	0.095	0.040	0.065	0.163	
	S.E.	(0.011)	(0.010)	(0.008)	(0.039)	
	N	661	425	994	92	1,086
85+	Mean	0.132	0.117	0.127	0.120	
	S.E.	(0.016)	(0.020)	(0.013)	(0.046)	
	N	477	256	683	50	733
Frequency		4,082	4,099	6,946	1,235	8,181

Table 2: Mortality Rates, Income and Smoking

at the age of sixty-five or when first observed, the lower are mortality rates. We treat self-reported health like smoking behaviour, as a control variable in our study.

3.2 Drivers of Income and Education

If income is an influence on mortality, there is a question what drives income. Educational attainment is universally regarded as an influence on income in working life. To the extent that income after retirement is influenced by income in working life, for example as a result of participation in pension schemes it is only to be expected that income after retirement is also influenced by educational attainment. There is a question how far this income represents double-counting of income earned during working life (because it arises from saving out of labour income), and how far it represents an additional benefit of educational attainment (because it arises out of employers' pension contributions which are not included in most measures of labour income). Table 4 shows a clear relationship between educational attainment and income at the age of sixty-five or when first observed if later. If someone educated to level 4 has studied for six more years than someone not educated beyond level 0, then the data imply a return per year of study of around eight per cent. The table shows clearly the impact of working on income. The figures for men educated to level 2, while elevated, do not appear to be distorted by any clear outlier.

Age		Initial Heath State					Freq.
		V. Poor	Poor	Fair	Good	Excellent	
65-69	Mean	0.085	0.043	0.029	0.005	0.016	
	S.E.	(0.041)	(0.014)	(0.007)	(0.002)	(0.006)	
	N	47	207	613	1,109	450	2,426
70-74	Mean	0.233	0.065	0.044	0.021	0.018	
	S.E.	(0.077)	(0.018)	(0.009)	(0.004)	(0.006)	
	N	30	186	519	1,069	444	2,248
75-79	Mean	0.412	0.085	0.056	0.048	0.015	
	S.E.	(0.119)	(0.029)	(0.012)	(0.007)	(0.006)	
	N	17	94	356	878	343	1,688
80-84	Mean	0.222	0.184	0.080	0.064	0.058	
	S.E.	(0.098)	(0.055)	(0.018)	(0.010)	(0.015)	
	N	18	49	225	551	243	1,086
85+	Mean	0.231	0.107	0.166	0.116	0.112	
	S.E.	(0.117)	(0.058)	(0.030)	(0.017)	(0.024)	
	N	13	28	151	362	179	733
Frequency		125	564	1,864	3,969	1,659	8,181

Table 3: Mortality Rates as a Function of Age and Initial Health State

Education		Respondent Works		Freq.
		No	Yes	
Level 0	Mean	9,647	14,798	
	S.E.	(257)	(1160)	
	N	598	54	652
Level 1	Mean	11,854	17,340	
	S.E.	(538)	(1299)	
	N	218	46	264
Level 2	Mean	13,266	28,839	
	S.E.	(794)	(4768)	
	N	99	12	111
Level 3	Mean	13,482	17,408	
	S.E.	(834)	(1578)	
	N	66	14	80
Level 4	Mean	16,409	23,414	
	S.E.	(747)	(2133)	
	N	114	39	153
Frequency		1,095	165	1,260

Table 4: Education and Initial Income

Educational attainment is often found to be related to parental occupation, income or social class. Table 5 shows that men with fathers in professional, managerial and technical occupations were much more likely to be educated to level 4 than were those whose fathers worked in partly-skilled or unskilled occupations. Conversely, the sons of fathers with unskilled or partly-skilled occupations were much less likely to be educated beyond level zero than were the sons of fathers with higher-grade occupations. These observations suggest that the role of father's occupation as an influence on educational attainment can be exploited when trying to understand the influence of educational attainment on income at the age of sixty-five.

3.3 Non-response

Finally, we need to consider the question of non-response. To the extent that non-response is correlated with the other variables of interest to us, it needs to be modelled separately. Table 6 provides a summary of non-response rates by age and initial health status. Not surprisingly, men in poor initial health are more likely to drop out of the survey, raising the possibility that men in very poor health and thus with low life expectancy are less likely to join the survey in the first place. It is not possible to investigate this.

Education		Father's Occupation				Freq.
		1. Professional/ Managerial	2. Clerical/ Craft/Service	3. Operative/ Other	Unclassified	
Level 0	Mean	0.378	0.456	0.663	0.550	
	S.E.	(0.032)	(0.023)	(0.024)	(0.039)	
	N	88	218	258	88	652
Level 1	Mean	0.146	0.259	0.180	0.225	
	S.E.	(0.023)	(0.020)	(0.019)	(0.033)	
	N	34	124	70	36	264
Level 2	Mean	0.150	0.094	0.046	0.081	
	S.E.	(0.023)	(0.013)	(0.011)	(0.022)	
	N	35	45	18	13	111
Level 3	Mean	0.094	0.069	0.044	0.050	
	S.E.	(0.019)	(0.012)	(0.010)	(0.017)	
	N	22	33	17	8	80
Level 4	Mean	0.232	0.121	0.067	0.094	
	S.E.	(0.028)	(0.015)	(0.013)	(0.023)	
	N	54	58	26	15	153
Frequency		233	478	389	160	1,260

Table 5: Education and Father's Occupation

Age		Initial Health State					Freq.
		Very Poor	Poor	Fair	Good	Excellent	
65-69	Mean	0.060	0.063	0.054	0.042	0.049	
	S.E.	(0.034)	(0.016)	(0.009)	(0.006)	(0.010)	
	N	50	221	648	1,158	473	2,550
70-74	Mean	0.063	0.046	0.039	0.023	0.031	
	S.E.	(0.043)	(0.015)	(0.008)	(0.005)	(0.008)	
	N	32	195	540	1,094	458	2,319
75-79	Mean	0.150	0.021	0.056	0.022	0.009	
	S.E.	(0.080)	(0.015)	(0.012)	(0.005)	(0.005)	
	N	20	96	377	898	346	1,737
80-84	Mean	0.053	0.039	0.022	0.038	0.024	
	S.E.	(0.051)	(0.027)	(0.010)	(0.008)	(0.010)	
	N	19	51	230	573	249	1,122
85+	Mean	0.000	0.176	0.074	0.062	0.032	
	S.E.	(0.000)	(0.065)	(0.020)	(0.012)	(0.013)	
	N	13	34	163	386	185	781
Frequency		134	597	1,958	4,109	1,711	8,509

Table 6: Non-response as a Function of Age and Initial Health

4 An Analysis of Educational Attainment, Initial Income, Survival and Response

Our approach to estimation is intended to reflect fully the inter-relationships between the different dependent variables of interest to us. However, because the educational attainment of the sample we study is inevitably influenced by factors rather different from those determining the other variables, we consider this separately from response, initial income and survival.

4.1 Educational Attainment

With the ranking of the educational categories shown in table 1, we explored an ordered probit equation as a means of explaining educational attainment. The effects of time are explained by a cubic polynomial[3] in year of birth, YB_i, measured relative to 1900. Three dummy variables for the three categories showing father's occupational status ($C_{i,1}, C_{i,2}$ and $C_{i,3)}$ (with zero values for all of them representing no answer to this question) are included with a further dummy ($D33_i$), indicating, as suggested by Silles (2009), whether the respondent was born before 1933. Such a dummy reflects the fact that the school leaving age was raised from fourteen to fifteen in 1947.

Formally, E_i denotes educational status ($E_i = 0$ to 4). We consider educational attainment to be determined by a latent variable e_i,

$$\begin{aligned} e_i &= \alpha_{01} YB_i + \alpha_{02} YB_i^2/100 + \alpha_{03} YB_i^3/1000 \\ &\quad + \alpha_{04} C_{i,1} + \alpha_{05} C_{i,2} + \alpha_{06} C_{i,3} + \alpha_{07} D33_i + \varepsilon_{0,i} \\ &= \mathbf{X}_i \boldsymbol{\gamma}_0 + \varepsilon_{0,i} \end{aligned} \tag{1}$$

where $E_i = k$ if $e_i > E_k^*$, with $E_0^* = -\infty$ and $\varepsilon_{0.i}\tilde{} N(0,1)$.

4.2 Initial Income, Survival and Response

We now move on to discuss the equations used to explain initial income, survival and response. It is helpful to set these out together, although not all variables are used in all equations. We define $S_{i,}$as a dummy which takes the value 1 if the respondent reported smoking when first observed and 0 otherwise, $H_{i,k}$ ($k = 1, 4$) as dummies indicating the

[3]We are grateful to a referee for suggesting this.

self-reported health status of the respondent on joining the panel (Very Poor to Good with Excellent omitted), and YJ_i, indicates the year in which the respondent joined the panel. t indicates the calendar year, $Age_{i,t-1}$ indicates the age of the respondent in the previous wave and $Ages_i$ the age of the respondent when first observed, W_i is a dummy which indicates whether the respondent is working when first observed, N_{it-1} is the number of waves to which the respondent has already replied, $ES_{i,k}$ (k=1,4) are dummies indicating highest educational attainment when first joining the sample with level 0 omitted, $Reg_{i,k}$ is a dummy which takes a value of 1 if the respondent lived in region k when joining the survey, $C_{i,1}$ is the dummy indicating that the respondent's father had a category 1 occupation and c is a constant term. Not all variables enter into all equations; in particular the equation for initial income is driven by age when observed and not by current age. Exclusion restrictions are discussed subsequently.

$$\mathbf{Z}_{it} = (S_{i,t}, H_{\ i,j}, YJ_i, t, Age_{i,t}, Ages_i, W_i, N_{it-1}, C_{i,1}, ES_{i,k}, Reg_{i,k}\ c)$$

The three further dependent variables are defined as follows. LY_i is the log of initial income on an equivalised basis. Q_{it} takes a value of 1 if the individual is reported alive and 0 otherwise, and R_{it} takes the value 1 if the individual responds to the survey but 0 otherwise. A record indicating that the respondent has died is regarded as a response.

The latent variables that underlie survival and response are q_i and $r_{i,}$ with $Q_{it} = 1$ if $q_{it} > 0$, and $R_{it} = 1$ if $r_{it} > 0$. The equations used were

$$LY_i = \mathbf{Z}_{it}\boldsymbol{\gamma}_1 + \varepsilon_{1,it} \tag{2}$$

$$q_{it} = \alpha_2 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_2 + \varepsilon_{2,it} \tag{3}$$

$$r_{it} = \alpha_3 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_3 + \varepsilon_{3,it} \tag{4}$$

4.3 Estimation: the System as a Whole

The covariance matrix of the residuals of the four interdependent equations, for educational status, income, survival and response, is given as

$$Cov \begin{bmatrix} \varepsilon_{0,i} \\ \varepsilon_{1,it} \\ \varepsilon_{2,it} \\ \varepsilon_{3,it} \end{bmatrix} = \begin{pmatrix} 1 & \sigma_{01} & \sigma_{02} & \sigma_{03} \\ \sigma_{01} & \sigma_{11}^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{02} & \sigma_{12} & 1 & \sigma_{23} \\ \sigma_{03} & \sigma_{13} & \sigma_{23} & 1 \end{pmatrix}$$

Rather than attempt to estimate a system in four variables, we estimate the education equation and introduce the generalised residuals from this into the other three equations.

This follows in the spirit of Kim (2004) who considers the case of a Markov switching model with an endogenous continuous regressor in the outcome equations. We make a Cholesky decomposition of the covariance matrix which maintains the correlation structure

$$\begin{bmatrix} \varepsilon_{0,i} \\ \varepsilon_{1,it} \\ \varepsilon_{2,it} \\ \varepsilon_{3,it} \end{bmatrix} = \begin{pmatrix} w_{11} & 0 & 0 & 0 \\ w_{12} & w_{22} & 0 & 0 \\ w_{13} & w_{23} & w_{33} & 0 \\ w_{14} & w_{24} & w_{34} & w_{44} \end{pmatrix} \begin{bmatrix} v_{0,i} \\ v_{1,i} \\ v_{2,it} \\ v_{3,it} \end{bmatrix}$$

Bringing together equations (1) to (4), the full model is

$$\begin{aligned} e_i &= \mathbf{X}_i\boldsymbol{\gamma}_0 + w_{00}v_{0,i} \\ LY_i &= \mathbf{Z}_{it}\boldsymbol{\gamma}_1 + w_{01}v_{0,i} + w_{11}v_{1,i} \\ q_{it} &= \alpha_2 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_2 + w_{02}v_{0,i} + w_{12}v_{1,i} + w_{22}v_{2,i} \\ r_{it} &= \alpha_2 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_3 + w_{03}v_{0,i} + w_{13}v_{1,i} + w_{23}v_{2,it} + w_{33}v_{3,it} \end{aligned}$$

This allows us to substitute for $v_{0,i}$ as

$$\begin{aligned} LY_i &= \mathbf{Z}_{it}\boldsymbol{\gamma}_1 + \frac{w_{01}}{w_{00}}\left(e_i - \mathbf{X}_i\boldsymbol{\gamma}_0\right) + w_{11}v_{1,i} \\ q_{it} &= \alpha_2 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_2 + \frac{w_{02}}{w_{00}}\left(e_i - \mathbf{X}_i\boldsymbol{\gamma}_0\right) + w_{12}v_{1,i} + w_{22}v_{2,i} \\ r_{it} &= \alpha_3 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_3 + \frac{w_{03w}}{w_{00}}\left(e_i - \mathbf{X}_i\boldsymbol{\gamma}_0\right) + w_{13}v_{1,i} + w_{23}v_{2,it} + w_{33}v_{3,it} \end{aligned}$$

Kim's approach addresses the case of a continuous endogenous regressor and involves including a residual term from the regression of the endogenous variable on instrumental variables uncorrelated with the error terms in the outcome equations in order to overcome the endogeneity-induced bias. The significance of the estimated coefficient attached to the residual term provides a test of endogeneity. Our case is slightly different in that the potentially endogenous regressor, educational status, is categorical rather than continuous. Following Vella & Verbeek (1999) and Orme (2001), we replace the term $(e_i - \mathbf{Z}_{it}\boldsymbol{\gamma}_0)$ with the generalised residual from equation (1), $\hat{\varepsilon}_{0,i}$. Since $\hat{\varepsilon}_{0,i}$ is not not correlated with $v_{1,i}$, $v_{2,it}$ or $v_{3,it}$, inclusion of this term, *Education Residual,* as a regressor in each of the other equations controls for the endogeneity of $ES_{i,k}$. Our model becomes:

$$\begin{aligned} LY_i &= \mathbf{Z}_{it}\gamma_1 + \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i} + w_{11}v_{1,i} \\ q_{it} &= \alpha_2 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_2 + \frac{w_{02}}{w_{00}}\hat{\varepsilon}_{0,i} + w_{12}v_{1,i} + w_{22}v_{2,it} \\ r_{it} &= \alpha_3 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_3 + \frac{w_{03w}}{w_{00}}\hat{\varepsilon}_{0,i} + w_{13}v_{1,i} + w_{23}v_{2,it} + w_{33}v_{3,it} \end{aligned}$$

By construction the v are independent normal variables with zero mean and unit variance. So the error terms can be written as $u_{1,i}, u_{2,it}, u_{3,it}$ with

$$Cov \begin{bmatrix} u_{1,i} \\ u_{2,it} \\ u_{3,it} \end{bmatrix} = \mathbf{CC}' \text{ with } \mathbf{C}= \begin{bmatrix} w_{11} & 0 & 0 \\ w_{12} & w_{22} & 0 \\ w_{13} & w_{23} & w_{33} \end{bmatrix}$$
$$= \begin{pmatrix} \sigma_{11}^2 & \sigma_{11}\rho_{12} & \sigma_{11}\rho_{13} \\ \sigma_{11}\rho_{13} & 1 & \rho_{23} \\ \sigma_{11}\rho_{13} & \rho_{23} & 1 \end{pmatrix}$$

and the identifying restrictions that $Var(u_{2,it}) = Var(u_{3,it}) = 1$. We estimate this system by maximum likelihood.

Consider the distributions of $u_{2,it}$ and $u_{3,it}$ conditional on the observed initial income.

$$f(u_{2,it}, u_{3,it}|u_{1,i}) \sim N\left\{ \begin{pmatrix} \alpha_2 LY_i + \mathbf{Z}_{it}\gamma_2 + \frac{\rho_{12}}{\sigma_1}\left(LY_i - \mathbf{Z}_{it}\gamma_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}\right) \\ \alpha_3 LY_i + \mathbf{Z}_{it}\gamma_3 + \frac{\rho_{13}}{\sigma_1}\left(LY_i - \mathbf{Z}_{it}\gamma_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}\right) \end{pmatrix}, \begin{pmatrix} 1-\rho_{12}^2 & \sigma_{23}-\rho_{12}\rho_{13} \\ \sigma_{23}-\rho_{12}\rho_{13} & 1-\rho_{13}^2 \end{pmatrix} \right\}$$

and set

$$\mathbf{V} = \begin{pmatrix} 1-\rho_{12}^2 & \sigma_{23}-\rho_{12}\rho_{13} \\ \sigma_{23}-\rho_{12}\rho_{13} & 1-\rho_{13}^2 \end{pmatrix}.$$

Bayes theorem indicates immediately that the likelihood of a given set of residuals, $u_{1,it}$, $u_{2,it}$ and $u_{3,it}$ is given as $f(u_{1,i})f(u_{2,it}, u_{3,it}|u_{1,i})$. Then, with $Q_{it} = 0$ and $R_{it} = 1$, i.e. if a death is reported, the likelihood function is, with $\phi()$ representing the density function of the normal distribution and $\Phi_2()$ representing the cumulative normalised bivariate normal distribution

$$L_{0,1,it} = \phi\left(\frac{LY_i - \mathbf{Z}_{it}\boldsymbol{\gamma}_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}}{\sigma_1}\right)\Phi_2\left\{-\frac{\alpha_2 LY_i + \mathbf{Z}_{it}\gamma_2 + \frac{\rho_{12}}{\sigma_1}\left(LY_i - \mathbf{Z}_{it}\boldsymbol{\gamma}_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}\right)}{\sqrt{1-\rho_{12}^2}},\right.$$
$$\left.\frac{\alpha_3 LY_i + \mathbf{Z}_{it}\gamma_3 + \frac{\rho_{13}}{\sigma_1}\left(LY_i - \mathbf{Z}_{it}\boldsymbol{\gamma} - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}\right)}{\sqrt{1-\rho_{13}^2}}, -\frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{13}^2}}\right\}$$

while with $Q_{it} = 1$ and $R_{it} = 1$

$$L_{1,1,it} = \phi\left(\frac{LY_i - \mathbf{Z}_{it}\boldsymbol{\gamma}_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}}{\sigma_1}\right)\Phi_2\left\{\frac{\alpha_2 LY_i + \mathbf{Z}_{it}\gamma_2 + \frac{\rho_{12}}{\sigma_1}\left(LY_i - \mathbf{Z}_{it}\boldsymbol{\gamma}_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}\right)}{\sqrt{1-\rho_{12}^2}},\right.$$
$$\left.\frac{\alpha_3 LY_i + \mathbf{Z}_{it}\gamma_3 + \frac{\rho_{13}}{\sigma_1}\left(LY_i - \mathbf{Z}_{it}\gamma_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}\right)}{\sqrt{1-\rho_{13}^2}}, \frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{13}^2}}\right\}$$

and if $R_{it} = 0$ and Q_{it} is not observed

$$L_{0,it} = \phi\left(\frac{LY_i - \mathbf{Z}_{it}\boldsymbol{\gamma}_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}}{\sigma_1}\right)\Phi\left\{-\frac{\alpha_3 LY_i + \mathbf{Z}_{it}\gamma_3 + \frac{\rho_{12}}{\sigma_1}\left(LY_i - \mathbf{Z}_{it}\boldsymbol{\gamma}_1 - \frac{w_{01}}{w_{00}}\hat{\varepsilon}_{0,i}\right)}{\sqrt{1-\rho_{13}^2}}\right\}.$$

The overall log likelihood function for observation t of individual i is then

$$L_{it} = L_{0,it}^{(1-R_{it})} L_{0,1,it}^{(1-Q_{it})R_{it}} L_{1,1,it}^{Q_{it}R_{it}}$$

The model was estimated using *STATA* version 13.

5 Results

5.1 Educational Attainment

Table 7 presents the parameters of the ordered probit model. In specifying this equation we require there to be instruments which allow us to address the possibility that joint unobserved factors influence both educational attainment and income at age sixty-five. Table 7 suggests a strong link between father's social class and educational attainment, with a father of high social standing raising the chance of high educational attainment, and a father of low social standing depressing the chance of high attainment. The unrestricted version includes a dummy for men born after 1933, and thus affected by the raising of the school leaving age in 1947. The results suggest that, while this is positive, it is a long way from significant; this dummy is excluded in the restricted version which we use in what follows.

As section 4.3 makes clear, the generalised residuals from these equations ($\hat{\varepsilon}_{0,i}$, *Education Residual*) are introduced as an explanatory variable in the subsequent models so as to address the possibility that effects which might otherwise be attributed to education are in fact explained by other influences which also affect educational attainment.

		Unrestricted			Restricted		
		Coef.	S.E		Coef.	S.E.	
Father's	Class 1	0.516	0.117	***	0.518	0.117	***
Occupational	Class 2	0.175	0.105	*	0.177	0.105	*
Status	Class 3	-0.317	0.111	***	-0.316	0.110	***
Year of Birth		0.237	0.110	**	0.237	0.110	**
Year of Birth^2/100		-0.700	0.340	**	-0.711	0.339	**
Year of Birth^3/1000		0.719	0.340	**	0.747	0.334	**
Post-1933 dummy		0.074	0.158				
	E_1	2.980	1.129	***	2.959	1.127	***
	E_2	3.584	1.130	***	3.564	1.128	***
	E_3	3.904	1.131	***	3.884	1.129	***
	E_4	4.199	1.132	***	4.179	1.129	***
Observations		1260			1260		
Pseudo R^2		0.0449			0.0448		
Log-Likelihood		-1580.66			-1580.77		
Significance Levels		*** 1%	** 5%		* 10%		

Table 7: Father's Background and Educational Attainment

5.2 Response, Survival and Income

Estimation of the remaining three equations leads to the set of parameters shown in table 8. Response is shown to be lower for smokers and negatively affected by income; it is, however, positively influenced by educational attainment. The coefficients on educational attainment in the income equation imply, nevertheless, that, on balance, response rates increase with educational attainment. Response rates are probably declining in initial health state; the coefficient for those in very poor health is less negative than for those in poor health, but the high standard error on the former means that the results are consistent with response rates being lower the worse is initial health. Men who are working when first observed are more likely to respond while response falls off with age. Other things being equal, the probability of dropping out of the survey decreases with the number of waves to which the respondent has already replied.

The survival equation shows clearly the effects which might be expected. Initial smoking behaviour, health status and income all influence subsequent survival significantly. It is likely that this equation is a reduced form summary of the influence of lagged income and health status on survival, because the values of these when first observed are likely to be strong predictors of their subsequent values. It is noticeable that whether someone works when first observed has very little direct influence on their survival rate, despite the fact that it has an influence on income when first observed. Since long-run income is presumably much less sensitive to employment status at the age of sixty-five than is income at this time, one might have expected a negative coefficient offsetting the fact that income is only temporarily elevated by employment. There are two obvious reasons why the effect of working when first observed should be fully represented by the income term. One is that differences persist beyond retirement, because men working later also receive larger retirement incomes; the other is that health benefits are conferred over and above those indicated by the health control variables. Both of these may be true to some extent.

Even with controls for health state, education and employment status, the income equation suggests that men who smoke at the age of sixty-five receive lower incomes than those who do not; the equation says nothing about the mechanism involved which is outside the scope of this study. The control variables indicating health state are all significant except for very poor health. While the difference between the coefficient on that and the coefficient on the dummy for poor health is probably not significant, it is quite likely that men in very poor health receive a range of benefits not available to those in better health, and this explains the difference between the two coefficients. The income equation shows, in broad terms, the sort of effects from education which would be expected in an equation explaining wages. Men educated to level 3 have incomes lower than those educated to level 2, but the different is not large and not statistically significant. The coefficient of 0.50 on level 4 education is broadly consistent with the findings of Dickson (2013) who suggested that each year of education raises wage income by about 10%. Men with level 4 qualifications are likely to have studied for about six years longer than those with no significant qualifications. This observation does not offer any verification of the coefficient because the factors, such as pension participation and the nature of pension arrangements, which influence income in old age are likely to be different from those which matter during working life. But it is striking nonetheless that the consequence of these is to preserve the sort of differentials which exist during working life. Working raises income at sixty-five or when first observed by 0.3 log units (35 per cent). As a post-script, we note that the generalised residuals from the education equation are not significant in any of the equations.

Table 9 shows the structure of the covariance matrix of the residuals in the multivariate system. The correlation between the unexplained components of initial income and response is highly significant, with the consequence that the hypothesis that all three residuals are uncorrelated is easily rejected ($\chi^2_6 = 193$). This suggests that the inter-dependent nature of the system is of statistical importance

		Response			Survival			Income		
		Coef	S.E.		Coef	S.E.		Coef	S.E.	
Smoke at start		-0.159	0.076	**	-0.252	0.071	***	-0.149	0.043	***
log Income at 65		-1.146	0.061	***	0.380	0.164	**			
Health	V. Poor	-0.179	0.200		-0.994	0.178	***	-0.045	0.095	
at 65	Poor	-0.328	0.117	***	-0.406	0.120	***	-0.165	0.071	**
	Fair	-0.222	0.090	**	-0.222	0.085	***	-0.098	0.048	**
	Good	-0.092	0.081		-0.009	0.075		-0.076	0.042	*
Highest	Level 1	0.381	0.125	***				0.190	0.068	***
Qual	Level 2	0.783	0.181	***				0.395	0.101	***
	Level 3	0.520	0.214	**				0.301	0.102	***
	Level 4	0.961	0.253	***				0.517	0.138	***
Works at start		0.721	0.114	***	0.042	0.109		0.308	0.056	***
Age		-0.023	0.005	***	-0.046	0.005	***			
Wave		-0.002	0.010		0.011	0.007				
London		0.087	0.135		0.147	0.119		0.114	0.086	
South-West		-0.075	0.145		0.215	0.116	*	-0.018	0.092	
East Anglia		0.182	0.163		0.131	0.132		0.060	0.094	
East Midlands		0.084	0.153		0.074	0.124		-0.010	0.085	
West Midlands		-0.172	0.152		0.212	0.128	*	-0.103	0.096	
North-West		-0.170	0.147		0.143	0.122		-0.014	0.091	
Yorks Humb.		0.051	0.152		0.217	0.129	*	-0.035	0.087	
North		-0.188	0.150		0.133	0.138		-0.049	0.094	
Wales		0.089	0.165		0.082	0.143		0.023	0.093	
Scotland		-0.098	0.155		0.104	0.128		0.018	0.093	
Waves already replied		0.066	0.013	***						
Age at Start								-0.009	0.003	***
Year of Start								0.012	0.005	**
Father Social Class 1								0.133	0.045	***
Education Residual		-0.139	0.099		-0.009	0.039		-0.057	0.054	
Constant		13.303	0.807	***	1.639	1.738		-13.398	9.729	
	Significance	10%	*	5%	**	1%	***			

Table 8: Determinants of Response, Survival and Initial Income

	Coef	S.E.	
$Tanh^{-1}\rho_{12}$	-0.350	0.733	
$Tanh^{-1}\rho_{13}$	0.612	0.047	***
$Tanh^{-1}\rho_{23}$	-0.121	0.083	
$\ln \sigma_{33}$	-0.736	0.035	***
Significance	Levels	10%	*
		5%	**
		1%	***

Table 9: Correlations and the Standard Error

The model set out in table 8 embodies a number of identifying exclusion restrictions. The selection equation is identified by a single variable, *Waves already replied,* which is highly significant. Its use as an instrument is justified by the observation that, to the extent that men's propensities to respond to the survey differ, those who have co-operated in the past are likely also to co-operate in the future. The role of income in the survival equation is identified by the assumption that education and father's social status, to the extent that they affect survival, do so through their influence on income. The non-linear nature of the system means that there is no means of testing this. But an indication of the validity of the assumption can be provided by treating survival as a continuous variable, and estimating the survival equation with income instrumented in the way it is in table 8. The Sargan test for over-identification is $\chi^2_5 = 0.2$. This gives no suggestion that the zero restrictions behind the specification of the survival equation are invalid. Finally it is necessary to comment on the fact that the dummy for fathers from social class 1 was included in the income equation, while the other social class dummies were not. Looking at the income equation on its own, inclusion of all the dummies results in a situation where it is identified by the quadratic and cubic terms in year of birth, and by the non-linear nature of the system. When estimated in this way, the other social class dummies were not significant, and they were therefore excluded.

We also looked at whether there was a role for the dummy for men born after 1933 in the income and survival equations, so as to explore whether our system might look for an effect from the raising of the school leaving age through a direct effect on earnings. The variable was not significant in either equation; in the income equation the coefficient was 0.1 with a z-statistic of 1. Thus our results appear compatible with the findings of Clark & Royer (2013) who reported an effect which was negative but not significantly different from zero.

6 A Reduced Form Alternative

The model set out above explores the effects of education subject to the restriction that their influence is felt entirely through income. In this section we explore a reduced form model in which the educational dummies enter directly into the survival equation, in place of the income term. We estimate a probit model with selection effects. The equations which explain educational attainment and response are as in section 5 and the generalised residuals from the former are introduced as explanatory variables in both the response and survival equations.

		Coef.	S.E.	
Highest	Level 1	0.174	0.122	
Qual	Level 2	0.241	0.169	
	Level 3	0.229	0.206	
	Level 4	0.498	0.241	**
Significance	Levels	10%	*	
		5%	**	
		1%	***	

Table 10: Education and Working Coefficients in a Reduced-form Model of Survival

The relevant coefficients are shown in table 10. That for level 4 education is significant at a 5 per cent level while the other coefficients are insignificant.

One interpretation of these coefficients is that only level 4 education affects mortality rates. But the ratios of the coefficients are not very different from those of the education terms in the income equation. We therefore construct a composite educational attainment variable, ED_i, which combines the attainment dummies using the coefficients in the income equation

$$ED_i = 0.180ES_{i,1} + 0.389ES_{i,2} + 0.292ES_{i,3} + 0.508ES_{i,4}$$

and use this in place of the individual educational dummies. This yields the coefficients shown in table 11. Here the restricted variable is significant at a 10 per cent level; the relevant restrictions are easily accepted ($\chi^2_3 = 1.8$, $p = 0.6$).It is perhaps not surprising that the composite variable is statistically less significant than was income in our earlier equation. That is what would be expected if the influence of education works through income.

	Coef.	S.E.	
Composite Educational Variable	0.718	0.396	*
Significance	Level	10%	*

Table 11: Educational and Working Coefficients in a Restricted Reduced Form Model of Survival

7 Model Simulation

Since the model is specified as a system of interrelated linear and probit equations, it is not possible to infer, simply by examining the coefficients, how education affects life expectancy. Simulation does, however, make it possible to examine the relationship between education and survival, and also allow us to reach some conclusions about the impact of education on discounted future income at the age of sixty-five. In turn this makes it possible to show the effect of education on an overall indicator of welfare.

Our model reduces to the form shown in equations (5 - 6). Here $\mathbf{Z}_{it}$ is the vector of exogenous variables (age, year, smoking status, health state and region when first observed), LY_i is the log of initial income and *Education Residual* is set to zero in the simulations.

$$LY_i = \mathbf{Z}_{it}\boldsymbol{\gamma}_1 + \varepsilon_{i,1} \tag{5}$$

$$q_{it} = \alpha_2 LY_i + \mathbf{Z}_{it}\boldsymbol{\gamma}_2 + \varepsilon_{it,2} \tag{6}$$

with

$$Cov\begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{it,2} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1 \\ \rho_{12}\sigma_1 & 1 \end{pmatrix} \tag{7}$$

To simulate the model we require appropriate values for $\mathbf{Z}_{it}$ and appropriate values for the relevant error terms. The latter are drawn from the bivariate distribution shown by equation (7). Conditional on some realisation of the error term in the first equation, denoted by $\hat{\varepsilon}_{i,1}$ and thus of initial log income, $\hat{LY_i}$,

$$\varepsilon^*_{it,2} = \frac{\rho_{12}}{\sigma_1}\hat{\varepsilon}_{i,1} + u_{it,2}, \qquad u_{it,2}\tilde{\ }N(0, 1-\rho_{21}^2).$$

so that the expected value of q_{it} is given as $\hat{q}_{it} = \alpha_2\hat{LY_i} + \mathbf{Z}_{it}\boldsymbol{\gamma}_2 + \frac{\rho_{12}}{\sigma_1}\hat{\varepsilon}_{i,1}$. The probability of survival conditional on any level of income and given the other exogenous variables is $\pi(\hat{q}_{it}) = \Phi\left(\hat{q}_{it}/\left(1-\rho_{21}^2\right)\right)$ where $\Phi()$ is the cumulative normal density function. The life expectancy of someone at age sixty-five E_i^{65} is then computed as the sum of the probabilities of surviving to any given age, with the latter in turn given by the product of the probability of surviving from year $t-1$ to year t :

$$E_i^{65} = \sum_{\tau=65}^{\tau_{\max}} \prod_{t=65}^{\tau} \pi(\hat{q}_{i\tau})$$

It is not practical to produce a model of income dynamics, rather than just initial income, as a part of this analysis of educational attainment, survival and response. Nevertheless, it is possible to say something about the path of income in later life based on our analysis of the determinants of initial income. Since the latter is explained by date of birth and age when initially observed, the assumption that the influence of age on income is independent of whether men participate in the survey or not allows us to use equation (5) to estimate the way in which expected income changes with age. For men whose initial incomes take the values fitted by equation (5) this in turn makes it possible to estimate the financial benefit conferred by education at any particular age. Combining this with a value of survival, that in turn makes it possible to provide an estimate of the benefits conferred by education after retirement. For the analysis to be consistent with the traditional analysis of the returns to education, we need to ensure that we are looking only at retirement income which is not result of past saving by the individual studied; that is already accounted for in wage income. The way in which we do this is discussed subsequently.

The expected return for someone who receives the expected income is, with $L\hat{Y}(\tau)$ expected income at age τ defined as above and δ a discount factor

$$HY_i^{65} = \sum_{\tau=65}^{\tau_{\max}} \exp\{L\hat{Y}(\tau)\}\delta^{\tau-65} \prod_{t=65}^{\tau} \pi(\hat{q}_{i\tau})$$

If we wish to value life itself, then, with V the value put on a life year, the discounted value of remaining life at the age of sixty-five is given as

$$DE_i^{65} = \sum_{\tau=65}^{\tau_{\max}} V\delta^{\tau-65} \prod_{t=65}^{\tau} \pi(\hat{q}_{i\tau}).$$

The choice of V is discussed subsequently.

7.1 Control Variables

In order to simulate the effects of education, it is necessary to choose appropriate values for the control variables. We replace the regional dummies by the proportions of the initial population in each region. We simulate for all combinations of smoking behaviour, health status, working status and parental social class. We then weight together the resulting simulation values by the proportions with which these characteristics are observed in the initial population. This means that, when we compare men with different educational attainment, we are doing it for a population which has the same average smoking, health, working and parental social status. Thus the simulations show the effects of education after controlling for these influences.

7.2 Parameter Uncertainty

The calculations above are performed for a fixed set of model parameters estimated as described above. But the standard errors associated with these parameters do not provide any direct indication of the uncertainty surrounding our estimates of the group averages of the variables of particular interest to us. These also have to be computed by simulation.

The procedure we use is to simulate the experiences of a population of fifty thousand men one thousand times, with random values for the model parameters redrawn for each of these thousand simulations from the distribution implied by their variance-covariance matrices and the assumption that they are jointly normally distributed. The relevant variance-covariance matrices are those associated with the parameters of tables 8 and 9. For any given set of model parameters,

$$\boldsymbol{\zeta} = [\boldsymbol{\gamma}_1, \alpha_2, \boldsymbol{\gamma}_2, \sigma_1^2, \rho_{12}],$$

we compute the mean values of aggregates of interest, $E^{65}(\boldsymbol{\zeta})$, $HY^{65}(\boldsymbol{\zeta})$ and $DE^{65}(\boldsymbol{\zeta})$. The means of these across the simulations provide estimates of the variables concerned. The standard errors of the simulations provide an indication of the reliability of the estimates.

In order to asses whether differences between aggregates for subpopulations, R and S are significant, it is necessary to take account of possible covariances between the disturbances to the two variables. This is most easily done by computing, for each simulation, the difference between the two aggregates, for example $E_R^{65}(\boldsymbol{\zeta}) - E_S^{65}(\boldsymbol{\zeta})$. The standard error of this can then be compared with its simulated mean so as to indicate whether $E_R^{65}(\boldsymbol{\zeta}) - E_S^{65}(\boldsymbol{\zeta})$ is likely to be of statistical significance. This allows us to estimate both the differences between income, life expectancy and the aggregated welfare measure for men of differing educational attainment and also provides standard errors of these estimates.

7.3 An Indicator of Welfare

It is necessary to put a value on life in order to value the impact of education on life expectancy. Mason et al. (2009) draw attention to a range of valuations between £30,000 and £70,000 at 2005 prices. The National Institute of Clinical Excellence used £ 30,000 at current prices in 2008 (National Institute of Clinical Excellence 2008, Chapter 8, p.54) to value healthy life, while Muller et al. (2011) use the much larger figure of £160,000 (US$265,000) in their study of the costs of pollution damage in the United States. Since, while we condition on existing health, we do not model men's health states over time. We adopt a value of £30,000 per year of life. This probably builds an element of caution into the results [4].

Secondly, as suggested earlier, some account needs to be taken of the fact that, to the extent that post-retirement income is a consequence of saving out of recorded labour income, the direct benefits of it have already been accounted for traditional estimates of the returns to education. Differences in income after age sixty-five are largely a consequence of differences in occupational pensions. These are financed both by employee contributions, which are included in conventional analysis of the returns to education, and employer contributions which are omitted. Thus the income differential needs to be multiplied by the ratio of employer contributions to total contributions in order to correct for this. The national accounts show that, on average employers contributed about 70% of the total cost of pensions[5] and we therefore used this ratio to compute the impact of education on post-retirement income over and above that accounted for by saving out of reported income accruing during working life. We therefore add the discounted value placed on extra life expectancy to 70% of the differential in discounted income to obtain an indicator of the welfare at age sixty-five associated with the different levels of educational attainment.

[4] An alternative approach to valuing life is provided by Murphy & Topel (2006). They base theirs on the utility enjoyed by people who are alive. But the practical problem with this approach is that it requires a cardinal utility function. The widely used CES function is negative unless some constant is added back on. The appropriate constant can be estimated only by forming a view about the level of consumption at which life becomes not worth living. Given the judgements involved it is not clear that the approach is superior to the methods surveyed by Mason et al. (2009)

[5] The average share of employer contributions in the total over the period 1974-1996 was 73%. Since 1997 the national accounts do not distinguish employee contributions from individual purchases of life insurance policies. Pensioners also typically received lump sums on retirement and we have implicitly assumed that these account for the large part of investment income received by those over sixty-five.

8 Post-retirement Benefits of Education

Using the methods described in section 7, we calculate the values of discounted life expectancy, discounted health-adjusted life expectancy and discounted income for men with each of the five levels of education which we identify. We also present estimates of the differences in these aggregates for someone educated to level 4 relative to someone educated to level 0.

The simulations generate an estimate for the average life expectancy of a man at age sixty-five of 18.3 years with a standard error of 1.0 years. The ONS estimate the expected remaining life of a sixty-five year old man in 1992 to be 16.2 years, rising to 18.0 years in 1997 and 20 years in 2006. Slightly more than half (681/1260) of our sample join at the start; these have an average age of 73.4. The life expectancy computed from this part of the sample should be expected to be below the cohort life expectancy at age sixty-five. But the remaining 579 are nearly all aged sixty-five when they join and should be expected to have the cohort life expectancy for sixty-five year olds in the year in which they join. So, while an exact comparison with official data is not possible, the estimated life expectancy for the sample is probably higher than that implied by the official data. Given the standard error it is probably not significantly higher. It is, however, quite likely that men in very poor health are less likely to participate in the survey in the first place (as opposed to dropping out after an initial response), and the parameter estimates show that very poor self-reported health has a substantial impact on mortality risk. It follows that the life expectancy of our sample should probably be expected to be higher than that shown in official data calculated from census records and registration of deaths.

Table 12 focuses on four variables calculated for a population with population average health, smoking and working status and regional mix, so that the differences arise only because of differences in the effects of educational status on initial income. First, simulated life expectancy is shown for each level of educational attainment. Then we present the discounted values of remaining income and discounted life expectancy. Finally, the table shows the estimates of initial income generated by the simulation model, for men aged sixty-five. These are, after allowing for the mix of men who are working and not working, very similar to the data means shown in table 4 of section 3.2. The table shows both the standard errors and means of the simulations. The results shown in this table are, as would be expected, fully consistent with the estimated parameters of the system. Those suggested that income at the age of sixty-five was significantly de-

		Level 0	Level 1	Level 2	Level 3	Level 4
Life	Mean	17.2	18.4	19.7	19.1	20.5
Expectancy	S.E.	(1.0)	(1.1)	(1.4)	(1.3)	(1.8)
Discounted	Mean	110,867	135,908	169,194	153,484	193,253
Income (£)	S.E.	(24833)	(32055)	(42553)	(39039)	(53876)
Discounted	Life	12.9	13.6	14.3	14.0	14.8
Expectancy	S.E.	(0.6)	(0.6)	(0.7)	(0.7)	(0.9)
Initial	Mean	10,373	12,560	15,472	14,086	17,514
Income (£)	S.E.	(1348)	(1628)	(2213)	(2060)	(2848)

Table 12: Life Expectancy and Income as Functions of Education

pendent on educational attainment and that survival depended significantly on income. The parameters suggested that level 3 education was worth less than level 2 education; this shows up in table 12.

In table 13 we show the differences between men educated to levels 1 to 4 and those educated only to level 0. The differences are calculated for populations which are in other respects similar. The differences in the means reflect the figures of table 12. Thus a man educated to level 4 can, at the age of sixty-five, expect 3.3 more years of life than can a man educated only to level 0. These differences are calculated for each of the thousand simulations, with the means and standard errors shown in the table; this method of calculation ensures that the standard errors and associated confidence intervals reflect the estimated covariance matrix of the parameters. The simulated confidence limits are provided by ranking the simulated differences and taking the 26th and 975th of the ranking. Comparing these limits with the standard errors suggests that the simulated values are not symmetric about the mean but are skewed to the right. All of the differences shown in the table are significant on a 95% basis. The relatively narrow confidence bands associated with the difference between level 0 and level 1 reflect the fact that, although the coefficient on the level 1 dummy in the income equation is smaller than that on the level 4 dummy, the standard error is also appreciably smaller.

Perhaps the most important feature of the table is the welfare indicator. This is calculated on the assumption that seventy per cent of the income difference between level 0 income and higher levels of income is not included in conventional measures of the return to education; neither of course is any of the direct benefit of increased longevity. As it turns out, increased longevity and increased income contribute to this in roughly equal proportions. Discounting at 3% p.a. back to the age of twenty-one, the benefit of a level 4 qualification relative to minimal qualifications is valued at just over

	Life Expectancy (years)	Discounted Income	Discounted Life Expectancy (years)	Initial Income	Welfare
			Level 1		
Mean	1.2	25,041	0.7	2,187	38,827
S.E.	0.7	11,539	0.4	835	19,075
Lower 95%	0.1	6,372	0.1	601	8,139
Upper 95%	2.9	51,166	1.7	3,959	82,587
			Level 2		
Mean	2.5	58,326	1.4	5,100	83,854
S.E.	1.3	23,288	0.7	1,586	32,607
Lower 95%	0.4	23,167	0.2	2,476	31,672
Upper 95%	5.4	113,729	3.1	8,600	158,830
			Level 3		
Mean	1.9	42,616	1.1	3,713	62,789
S.E.	1.1	20,235	0.6	1,448	29,193
Lower 95%	0.3	12,849	0.2	1,279	16,803
Upper 95%	4.4	88,747	2.5	6,914	131,214
			Level 4		
Mean	3.3	82,386	1.9	7,141	114,172
S.E.	1.8	35,927	1.0	2,399	48,391
Lower 95%	0.5	28,823	0.3	2,987	37,420
Upper 95%	7.2	166,741	4.1	12,305	220,300

Table 13: Effects of Educational Attainment relative to Level 0

£30,000; this gives some idea of the value to a young man of the benefits omitted from conventional estimates of the return to education.

9 Conclusions

In this paper we have explored the relationship between income, health, mortality and education in the population aged sixty-five and over. The analysis suggests a clear link between income when first observed and survival, with education affecting income. Our figures put a value of higher education relative minimal education discounted to age twenty-one of just over £30,000, with smaller sums for lower levels of educational attainment. These findings are broadly coherent with the differences in life expectancy at sixty-five by social class. They are probably somewhat weaker than those suggested by Lleras-Muney (2005). If someone educated to level 4 has studied for six years longer than someone with no qualifications and enjoys an expected 3.3 years of extra life, that implies an effect of 0.6 years of expected extra life per year of study although it has to be remembered that this is for someone aged sixty-five while the figure of 1.7 years per year of study quoted by Lleras-Muney (2005) is for a thirty-five year old. They are, nevertheless, at odds with the results presented by Clark & Royer (2013) for the impact of the change in the school leaving age in 1947 and showing if anything a reduction in survival rates. Of course they study the population aged forty-five to sixty-nine while we focus on men aged sixty-five and over. But there is another and more plausible explanation of this apparent difference.

The route by which education influences survival in our paper is through its influence on income at age sixty-five as a proxy for income then and in later years; we have shown that the reduced form analogue, while being less well determined than the structural model, is entirely consistent with it. Oreopoulos (2006) suggests that the increase in years of education had an impact on earnings resulting from the change in the school-leaving age comparable with what would be expected on the basis of earlier studies of returns to education, and also the more recent work of Dickson (2013). But that does not mean that there was the same significant differential effect on income at age sixty-five. We found that the raising of the school leaving age in 1947 had only an insignificant effect on income, while our educational attainment terms are highly significant and larger.

Recently many employers have reduced their pension commitments. Nevertheless Forth & Stokes (2010) show that private sector employers continue to make contributions. The mean contribution depends on the nature of the scheme but, for those with defined contribution arrangements, it found that the mean employer contribution to defined contribution occupational schemes was 14% of pay while to personal pension schemes it was 9% of pay. Thus, despite the general perception of widespread reductions in employer contributions, they remain substantial. Unless they fall further, they will continue to comprise an important component of the return to education omitted from conventional analysis.

References

Baker, D., Leon, J., E.G.Smith Greenaway, Collins, J. & Movit, M. (2011), 'The Education Effect on Population Health: a Reassessment', *Population Development Review* **37**, 307–332.

Barker, D., Eriksson, J., Forsen, T. & Osmond, C. (2002), 'Foetal Origins of Adult Disease: Strength of Effects and Biological Basis', *International Journal of Epidemiology* **31**, 1235–1239.

Batty, G., Derr, G., Macintyre, S. & Deary, I. (2006), 'Does IQ explain Socioeconomic Inequalities in Health? Evidence from a Population-based Cohort Study in the West of Scotland', *British Medical Journal* . doi:10:1136/bmj.38723.660637.AE.

Case, A. & Paxson, C. (2011), 'The Long Reach of Childhood Health and Circumstance: Evidence from the Whitehall II Study', *Economic Journal* **121**, F183–F204.

Clark, D. & Royer, H. (2013), 'The Effect of Education on Adult Mortality and Health: Evidence from Britain', *American Economic Review* **103**, 2087–2120.

Contoyannis, P., Jones, A. & Rice, N. (2004), 'The Dynamics of Health in the British Household Panel Survey', *Journal of Applied Economics* **19**, 473–503.

Dickson, M. (2013), 'The Causal Effects of Education on Wages Revisited', *Oxford Bulletin of Economics and Statistics* **75**, 477–498. DOI: 10.1111/j.1468-0084.2012.00708.x.

Economou, A. & Theodossiou, I. (2011), 'Poor and Sick: Estimating the Relationship between Household Income and Health', *Review of Income and Wealth* **57**, 395–411.

Eide, E. & Showalter, M. (2011), 'Estimating the Relation between Health and Education: what do we know and what do we need to know?', *Economics of Education Review* **30**, 778–791.

Forth, J. & Stokes, L. (2010), Employers' pension provision survey 2009. Department of Work and Pensions Research Report 687 http://research.dwp.gov.uk/asd/asd5/rports2009-2010/rrep687.pdf.

Frijters, P., Haisken-DeNew, J. & Shields, M. (2005), 'The Causal Effect of Income on Health: Evidence from German Re-unification', *Journal of Health Economics* **24**, 997–1017.

Fujiwara, T. & Kawachi, I. (2009), 'Is Education Causally Related to Better Health? A Twin Fixed-effect Study in the USA', *International Journal of Epidemiology* **38**, 1310–1322.

Gould, E., Lavy, V. & Paserman, M. (2011), 'Sixty Years after the Magic Carpet Ride: the Long-run Effect of the Early Childhood Environment on Social and Economic Outcomes', *Review of Economic Studies* **78**, 938–973.

Jenkins, S. (2010), The British Household Panel Survey and Its Income Data. Institute for Social and Economic Research, University of Essex Working paper No 2010-33.

Khoman, E., Mitchell, J. & Weale, M. (2008), 'Incidence-based estiamtes of life expectancy of the healthy for the uk: Coherence between transition probabilities and aggregate life tables', *Journal of the Royal Statistical Society: Series A* **171**, 203–222.

Kim, C. (2004), 'Markov-switching models with endogenous explanatory variables', *Journal of Econometrics* **122**, 127–136.

Lager, A., Bremberg, S. & Vagero, D. (2009), 'The Association of Early IQ and Education with Mortality: a 65-year Longitudinal Study in Malmo, Sweden', *British Medical Journal* **339**, b5282. doi:10.1136/bmj/b5282.

Levy, H. & Jenkins, S. (2008), Bristish Household Panel Survey Dervied Current and Annual Net Household Income Variables, Wave 1-16, 1991-2007. UK Data Archive Study Number 3909.

Lindahl, M. (2005), 'Estimating the Effect of Income on Health and Mortality using Lottery Prizes as an Exogenous Source of Variation in Income', *Journal of Human Resources* **40**, 144–168.

Lleras-Muney, A. (2005), 'The relationship between education and adult mortality in the united states', *Review of Economic Studies* **72**, 189–221.

Madsen, M., Andersen, A.-M., Christensen, K., Andersen, P. & Osler, M. (2009), 'Does Educational Status Impact Adult Mortality in Denmark? A Twin Approach', *American Journal of Epidemiology* **172**, 225–234.

Marmot Review (2010), Fair Society, Healthy Lives. Report of the Strategic Review of Health Inequlaities in England post-2010.

Mason, H., Jones-Lee, M. & Donaldson, C. (2009), 'Modelling the Money Value of a QALY: a New Approach based on UK Data', *Health Economics* **18**, 933–950.

Muller, N., Mendelsohn, R. & Nordhaus, W. (2011), 'Environment Accounting for Pollution in the United States Economy', *American Economic Review* **101**, 1649–1675.

Murphy, K. & Topel, R. (2006), 'The Value of Health and Longevity', *Journal of Political Economy* **114**, 871–904.

National Institute of Clinical Excellence (2008), The Guidelines Manual. http://www.nic.org.uk/niceMedia/pdf/GuidelinesManualChapter8.pdf.

Oreopoulos, P. (2006), 'Estimating average and local average treatment effects of education when compulsory schooling laws really matter', *American Economic Review* **96**, 152–175.

Orme, C. (2001), Two-step Inference in Dynamic Non-linear Panel Data Models. http://personalpages.manchester.ac.uk/staff/chris.orme.

Silles, M. (2009), 'The Causal Effect of Education on Health: Evidence from the United Kingdom', *Economics of Education Review* **28**, 122–128.

Smith, J. (1999), 'Healthy Bodies and Thick Wallets: the Dual Relation between Health and Economic Status', *Journal of Economic Perspectives* **13**, 145–166.

Vella, F. & Verbeek, M. (1999), 'Two-step estimation of panel data models with censored endogenous variables and selection bias', *Journal of Econometrics* **90**, 239–263.

Whalley, L. & Deary, I. (2001), 'Longitudinal Cohort Study of Childhood IQ and Survival up to Age 76', *British Medical Journal* **322**, 1–5.