National Research and Development Centre
for adult literacy and numeracy

Education, basic skills and health-related outcomes

Research report

Augustin de Coulon, Elena Meschi and Marisa Yates

Leading education and social research
Institute of Education
University of London

Published by the National Research and Development Centre
for Adult Literacy and Numeracy

For information on alternative formats, or to give feedback on the content and accessibility of this publication, please contact:

Publications
NRDC
Institute of Education
20 Bedford Way
London WC1H 0AL
Telephone: +44 (0)20 7612 6476
Fax: +44 (0)20 7612 6671
Email: publications@nrdc.org.uk

Contents

Preface: Policy context

Since this research was undertaken, the government has committed itself to the ambition of being a world leader in skills by 2020, benchmarked against the upper quartile of OECD[1] countries (Leitch 2006). For Skills for Life, this means ensuring that, by 2020, 95 per cent of adults possess at least functional levels of literacy and numeracy – defined as Entry level 3 numeracy and Level 1 literacy.

To make progress towards this ambition, the government has a Public Service Agreement (PSA) target that between 2008 and 2011:

- 597,000 people of working age achieve a first Level 1 or above literacy qualification;
- 390,000 people of working age achieve a first Entry level 3 or above numeracy qualification.

To deliver this target and in doing so drive progress towards the 2020 ambition, the government published a refreshed Skills for Life strategy in March 2009 (DIUS 2009). The refreshed strategy focuses on three central themes:

1. Focusing Skills for Life on employability, ensuring that the literacy, language and numeracy skills we help people develop will support them to find, stay and progress in work.
2. Raising demand for literacy, language and particularly numeracy skills among individuals and employers, changing the culture and attitudes to Skills for Life that prevent people from embarking on learning.
3. Delivering flexible and responsive Skills for Life provision which meets learner and employer needs, is high quality, delivered in innovative ways and embedded in wider skills provision where that is the best way to meet individual learners' needs.

The priority learner groups identified within the overall refreshed strategy are:

- people who are unemployed and on benefits;
- low-skilled adults in employment;
- offenders in custody and those supervised in the community; and
- other groups at risk of social exclusion.

Since this research pre-dates the refreshed Skills for Life strategy, there may be information in this document that relates solely to the original strategy and information which does not reflect more recent developments including those set out in the refreshed Skills for Life strategy.

[1] Organisation for Economic Co-operation and Development.

1. Introduction

The UK's low record in numeracy, literacy and ICT skills among the adult population has been widely documented (see for example the Moser Report (DfEE 1999) and the Leitch Review of Skills (2006)). Different studies have shown that low basic skills are associated with poor labour market outcomes (see for example de Coulon et al. 2007) and with children's lower cognitive performance (de Coulon et al. 2008).

This report investigates the effects of individuals' skills on wider outcomes related to health, particularly behaviours towards smoking, drinking, and body weight. All three behaviours have an impact on individual health: long-term use of cigarettes and alcohol have well-documented physical effects, such as pulmonary and liver diseases, and being overweight, particularly being obese, has been linked to high blood pressure and high cholesterol, as well as inducing increased risk of diabetes. The main aim of the report is to investigate how better skills (measured by basic skills tests' scores and highest qualification attained) are associated with a healthier lifestyle. One novelty of this report lies in the possibility to explore the differential role of formal education and actual basic skills as assessed by literacy and numeracy tests. Previous research has mainly focused on one of these variables only, failing to individuate the possible cumulative and interactive role of education and basic skills in affecting health behaviours. Moreover, we extend previous analysis by investigating not only the link between human capital and the *occurrence* of health-risky behaviours but also by exploring the link between human capital and the *amount* of alcohol and cigarettes consumed. Finally we also investigate how human capital is associated with changing behaviours towards health over the age range 16 to 34.

The reminder of the paper is organised as follows. Section 2 reviews the literature documenting the positive correlation between skills and health and provides an overview of the main mechanisms put forth in the literature to explain this positive relationship. In particular, three mechanisms are described:

1. Individuals with higher reading skills are better able to understand media coverage of new results from preventative health literature as well as government health awareness campaigns.
2. More literate individuals can also better interpret and follow medical prescription directions.
3. Higher educated individuals may have lower time preference, making them not only more likely to invest in education but to also follow a healthier lifestyle.

In Section 3 we discuss our empirical strategy, while in Section 4 we describe the data and provide the relevant descriptive statistics. We make use of the British Cohort Survey 1970 (BCS70), focusing in particular on the most recent sweep collected in 2004 where 9000 individuals were tested on their numeracy and literacy skills. The dataset is very rich, which allowed us to introduce a wide range of control variables together with the highest level of qualification attained. It was also possible to use the longitudinal properties of the BCS70, as individuals were interviewed on the same health-related outcomes in previous sweeps. Our results are presented and discussed in Section 5. The final section provides a summary of the main results and draws some policy implications as well as directions for future research.

2. Literature review

There is a large amount of literature documenting the strong correlation between education and health outcomes. Since the seminal work of Grossman (1972), many studies have tried to analyse the impact of education on different health outcomes. Grossman (1972) developed a theoretical model based on the argument that critical and analytical skills acquired by individuals during their education process are subsequently employed to improve efficiency of health production[2]. Several subsequent works have found evidence in support of a productive efficiency of schooling (see for example Wagstaff 1986, Erbsland et al. 1995 and Gilliskie and Harrison 1998)[3].

There are different mechanisms through which education may lead to better health. First of all, education provides individuals with better access to information and improved critical skills which makes them more conscious about the health consequences of risky behaviour (Wolf et al. 2007). A more educated person may also have a better understanding of his or her symptoms and, thus, be better able to explain to a doctor what they are, resulting in more effective treatment and better health outcomes later in life (Adams 2002). Moreover, education may improve health through increased social standing[4]: higher educated individuals often obtain better jobs that offer safer working environments, provide better health insurance and benefits, etc.

Empirically, the recent literature agrees on the existence of a robust positive association between educational levels and health, and this relationship has been observed across many countries and time periods, and for a wide variety of health measures (see Leigh 1998, Kenkel et al. 2006, Culter and Lleras-Muney 2006, Groot and Maassen van den Brink 2007).

However, it is not clear whether this correlation truly reflects a causal link or if it arises from unobserved 'third' factors that affect both educational attainment and health. That is, instead of being linked to improved health directly, educational attainment may also be an outcome of factors that also enhance health. For example, either genetic endowments or social background may jointly impact schooling and health outcomes. Richer families are more likely to invest more in their children's education and health; smarter individuals may be more likely to obtain more schooling and also take better care of themselves (Culter and Lleras-Muney 2006). Furthermore, factors such as time preference and self-efficacy may impact on both education and health levels. Individuals more willing to delay gratification (lower rates of time discount) are more likely to stay in school longer and do things that contribute to better health, such as eating healthily and

[2] Grossman modelled health outcomes using a production function framework where health-orientated behaviour is a process of combining a set of inputs in accordance with the state of technical knowledge in order to produce health stock. Health is viewed as a durable capital good that is desirable due to the fact that it produces 'healthy time'. Individuals are endowed with an initial stock of health that depreciates with age and is augmented through investment. The optimal level of health stock is at the point where the incremental gain from an additional unit of investment equals the cost it involves. Higher education levels induce the demand curve for health to shift outwards thereby increasing the demanded optimal health stock for given levels of health inputs. This is because higher education increases the efficiency with which investments in health are made.

[3] For a careful review of these studies see Grossman and Kaestner (1997) and Grossman (2005).

[4] See for example Marmot (2004). In a multi-decade study of the British civil service, he found that occupational status hugely affected health outcomes – independent of every other factor. He showed that it is not only the status that gives access to more information and a more health-safe environment, but that status itself played a key role, even if everything else was the same.

exercising – all behaviours which do not have immediate results or rewards. Increased self-efficacy – an individual's control over his or her behaviour – can also lead to increases in time spent in education as well as engaging in more health-positive behaviours (Fuchs 2004). Groot and Maassen van den Brink (2007) also argue that there may be an endogeneity problem in the sense that a higher education not only leads to a healthier lifestyle and better health, but also that people who are healthier are more efficient learners of new skills and competencies and invest more in education.

However, it seems that these 'third' factors do not fully explain the observed relationship. Several empirical papers show that even after including in the analyses measures of family background and detailed individual characteristics, the effect of education remains positive and significant, although lower in magnitude (e.g. Culter and Lleras-Muney, 2006, Kenkel et al. 2006). Other attempts to uncover a causal relationship have been achieved using an Instrumental Variables (IV) approach. For example, Adams (2002) uses the quarters of birth[5] to identify an exogenous variation in education and concludes that the education effect on health (measured using self-reported functional ability) is independent of the influence of omitted variables, as the education coefficient remains positive and significant in the IV estimates. Groot and Maassen van den Brink (2007) also find a positive and significant impact of education on self-reported health status, using Ordinary Least Squares (OLS), ordered probits and IV methods[6] (the two instruments for education are whether the father had a managerial job and the number of workers supervised by the father when the respondent was 14 years old, and whether the mother of the respondent (ever) had a paid job).

Kenkel et al. (2006) explore the relationship between high school completion and two possible causes of death: smoking and obesity. Using state educational policy measures as instruments for education, they find higher schooling significantly reduces the probability of being a smoker, while no significant impact emerged on the probability of being overweight or obese. Sander (1995) estimates IV models of the impact of schooling on smoking and quitting smoking. His instruments include parents' schooling, number of siblings and region of residence at age 16 and his results show that schooling increases the probability that men and women quit smoking. In a further study (Sander 1998), Sander shows that not only schooling, but also mental ability and time preference (future education) affect the probability of smoking and the amount of cigarettes smoked per day. No significant relationship is found however between college attendance and marijuana use. In another paper, Sander (1999) finds that educational attainment is inversely related to heavy drinking. Interestingly, he stresses that an important implication of his results is that the relationship between education and health habits might be specific to the habit in question and suggests caution in generalising the relationship.

[5] Quarter of birth are supposed to affect one's educational attainment due to the way in which compulsory school laws operate in the US (see also Angrist and Krueger 1991).

[6] They stress the importance of assessing the causality of the education-health link especially when measures of self-reported health are used, as "it is possible that higher educated people answer questions on their health status in a different way than lower educated people. Notions on what constitutes a good or a bad health may differ between higher and lower educating people, resulting in a spurious education effect on health. This phenomenon is known in the literature as scale of reference bias" (Groot and Maassen van den Brink, 2007, p 187.)

Aside from this literature, there is a growing body of medical evidence linking health literacy and health outcomes (see DeWalt et al. 2004 for a review and Kripalani et al. 2006). Amongst health outcomes investigated are depression (Weiss et al. 1992), asthma (Mancuso and Rincon 2006), HIV (Osborn et al. 2007), epilepsy (Shetty et al. 2007) and glaucoma (Muir 2006). Most of this literature focuses on how literacy may impact various aspects of medication use and adherence. The Test of Functional Health Literacy in Adults (TOFHLA) is often used as it has 17 items assessing individuals' capacity to read and understand hospital documents and labelled prescription vials. The test is highly correlated with other standard literacy tests (Baker 2006).

In another contribution, Chandola et al. (2006) have modelled different pathways between education and health condition using a structural equations modelling approach. They found that most of the link between education and adult health is explained by the mediation of adolescent health, health behaviours and the sense of control.

Sentell and Halpin (2006) examined the relation between literacy and health disparities and found that higher levels of literacy and higher educational levels were associated with better health. The researchers showed that the addition of literacy into their health models removed the predictive power of race and education and increased the overall predictive power of the model. This suggests that literacy contributes to the differences in overall health between individuals, above and beyond any contributions made by their educational attainment and their ethnic identity – two factors that are often associated with health differences in the general population. In general, the impact of basic skills on health tends to relate more highly to behaviours and outcomes that can be controlled (such as smoking, drinking or general overall health; DeWalt et al. 2004).

3. Empirical strategy

This section presents and discusses our empirical strategy to estimate the impact of education on different health outcomes.

As mentioned in the previous section, the literature has suggested different pathways through which education may affect health-related behaviour and outcomes. On the one hand, there could be a direct causal link due to health information: having better skills should help individuals in acquiring and understanding more information about the health consequences of risky behaviours. On the other hand, the relationship between education and health may be indirect, reflecting some unobserved characteristics that differ according to individuals' educational levels. In other words, more educated people may have some unobserved traits (such as family background, genetic traits or other individual differences such as the ability to delay gratification) that are also related to health outcomes, which makes the health-education link a non causal one.

Therefore the main aim of our empirical investigation is to understand whether individuals' skills (measured both by basic skills assessments and highest education achieved) and different health outcomes in adulthood are associated. We first focus on health outcomes at age 34 and test whether they are related to individuals' measures of skills. Given the richness of our dataset, we are able to control for a very large array of potential confounding factors.

We then try to identify a closer causal relationship by exploiting the longitudinal nature of our data and using an estimation method that takes into account the unobserved individual (time constant) characteristics.

The empirical analysis makes use of a very rich longitudinal database, the 1970 British Cohort Study (BCS70), which is based on a cohort of 17,196 babies born in England, Wales, Scotland and Northern Ireland[7] between the 5th and the 11th of April in 1970. Since the start of the BCS70, there have been seven waves of survey data taken: at age 5 (1975), at age 10 (1980), at age 16 (1986), at age 21 (a small subset of the original sample; 1991), at age 26 (1996), at age 30 (2000) and at age 34 (2004), and at every wave, questions relating to each cohort member's health, education, social and economic circumstances have been ascertained, normally via in-person interview.

3.1 Analysis at age 34

To test the relationship between education and health at age 34 we estimate the following regression:

$$H_i = \alpha + \beta HC_i + \gamma_k \sum_k X_{ik} + \varepsilon_i \qquad (1)$$

[7] After the initial survey in 1970, the cohort members in Northern Ireland were dropped; hence, the cohort is comprised of individuals from Great Britain only.

Where the subscript i is the individual indicator; H is a measure of health-related outcome; HC is a proxy of individuals human capital (see below); X_k are a set of control variables including individual characteristics and family background; finally ε is the usual error term, independently and identically distributed. We estimate different version of equation (1) corresponding to four health behaviours: whether smoking, whether being binge drinking or having some drinking problems, and whether being overweight or obese. We will describe these variables in more details in the next section where we also present some descriptive statistics.

We will focus our attention on β which measures the effect of education or basic skills on the particular measure of health behaviour. In our analysis we will measure human capital by considering both formal education and basic skills. The former is measured including dummy variables describing the highest qualification achieved. The latter are measured using a synthetic index created employing a Principal Component Analysis (PCA) based on literacy and numeracy test scores (details on the construction of this indicator are given in the next section).

The joint inclusion of a measure of basic skills beside the standard measure of formal education achieved constitutes an important novelty of this report. There are a number of reasons why we think that the inclusion of an indicator of basic skills constitutes an improvement with respect to previous works. First, basic skills and formal education – although related and sometime overlapping concepts – are in fact capturing different aspect of individuals' human capital as we will discuss in the next section. It is therefore important to analyse the role of these different aspects in health outcomes. In particular, we think it is especially important to include literacy and numeracy levels as they are potentially better measures of how efficient individuals are at producing health-related outcomes: health knowledge is probably more highly correlated with basic skills than with formal education levels. For example, low health literacy is associated with less medical knowledge, infrequent receipt of preventive services, increased hospitalisation and use of emergency care, and worse control of chronic diseases (Kripalani et al. 2006). Moreover, literacy and numeracy tests are designed to discriminate people at the bottom of the skill distribution with a low level of numeracy and literacy. This is particularly interesting in this context since health returns to education are shown to be larger for persons with low levels of education and skills (Grossman, 2004). For example, Kenkel et al. (2006, p. 12) stress that the informational advantage associated with schooling past high school is minimal. In this sense, adding an indicator of basic skills beside the formal education one, we are more able to identify variation among low-educated people.

As far as mediating variables are concerned, the use of a rich longitudinal data allows us to control for a large set of individual characteristics, socio-economic background and early family environment as well as different psychologically-related variables. The inclusion of such a vast array of explanatory variables reduces the probability that the 'omitted variables' problem may bias our coefficients.

In particular, we will include personal characteristics such as sex, marital status and employment status (whether full-time employed or not); socio-economic characteristics such as social class (described using the NS-SEC[8] occupationally-based classification) and income (this is measured as the log of the cohort member's weekly income). We are therefore controlling for the availability of financial resources which may affect access to health care, and, for the types of occupation: it may in fact be that 'better" jobs also offer safer work environments. The coefficient of the basic skills variable should therefore be interpreted as the incremental impact of basic skills conditional on the level of income and within any particular category of job. We also control for the number of children which may affect and change individuals' attitudes toward health-damaging behaviours. We then control for the different access to information by including a variable reflecting the frequency by which individuals usually read newspapers. We also include a variable describing the cohort members' evaluation about the quality of their local health services. This variable is aimed at capturing differential access to information and to health services which may impact on the awareness of the health consequences of risky behaviours. Finally, we include some psychological variables that may affect the propensity to engage in health-risky behaviours. We therefore insert an index of locus of control – which refers to the extent to which individuals believe that they can control events that affect them – and a measure of life satisfaction. The former is meant to capture the extent of individuals' self-efficacy which may lead them to engage in more health-positive behaviours (see Fuchs 2004), while the latter should proxy individuals' happiness and satisfaction that can also affect people's attitude with respect to their health.

With regard to the econometric method, the modelling of a dichotomous variable (in this case smoke/do not smoke; binge drinker/not a binge drinker; heavy drinker/not a heavy drinker; overweight/not overweight; obese /not obese[9]) may be undertaken using a probit model which runs as follows. Let Y be a binary outcome variable, and let X be a vector of regressors. The probit model assumes that:

$$\Pr\left(Y = 1/X = x\right) = \Phi(x'\beta) \qquad (2)$$

where Φ is the cumulative distribution function of the standard normal distribution. The parameters β are typically estimated by maximum likelihood. The probit model can be generated by a simple latent variable model. In other words, the probit model postulates the existence of a latent, unobserved, variable, Y_i*, which is related to a set of explanatory variables by the following relationship:

$$Y_i* = \beta' X_i + \varepsilon_i \qquad (3)$$

where ε_i is a random error term assumed to be normal. Although this latent variable is not observed, it is regarded as determining the value of the observed binary outcomes in this way:

$$\begin{aligned} Y_i &= 1 \quad if \quad Y_i* > 0 \\ Y_i &= 0 \quad otherwise \end{aligned} \qquad (4)$$

[8] National Statistics Socio-Economic Classification.
[9] See the next section for a detailed definition of these variables.

We will use these probit estimates to identify the impact of education and basic skills on the probability of being a smoker, a heavy/binge drinker and obese. We will then study the impact of the same variables (basic skills and education) on cigarette and alcohol consumption (looking at number of cigarettes smoked per day and daily alcohol consumption). When modelling the consumption of tobacco or alcohol (differently from other basic goods), it is important to take into account the high percentages of zeros in the data. Such non-observable values are non-random but are the results of the choices people make. In other words, individuals must pass two hurdles before being observed with a positive level of consumption. Both hurdles are outcomes of individual choices: a participation decision and a consumption decision (see Jones 1989).

We first use a Tobit model to take into account that the data are censored at 0. The Tobit estimator is based on two important pieces of information for each individual: the probability that an individual's score on the dependent variable is above the censoring threshold and the density of the dependent variable given that an individual scores above the censoring threshold. By explicitly incorporating both pieces of information into the likelihood function, the Tobit estimator provides consistent estimates of parameters governing the distribution of a censored normal random outcome variable (Smith and Brame 2003).

It is also possible to directly investigate the two individual choices (participation and consumption) by using a Heckman selection model[10]. The approach relies on the estimation of two equations: the outcome equation and the selection equation (whether to smoke or not). The outcome equation (consumption) can be expressed by a simple linear model:

$$Y_i = x_i'\beta + \varepsilon_i \qquad \textit{consumption equation}$$

where Yi is observed if a second, unobserved latent variable exceeds a particular threshold.

$$z_i^* = w_i\alpha + u_i \qquad z_i = \begin{cases} 1 & if \quad z_i^* > 0 \\ 0 & otherwise \end{cases} \qquad \textit{selection equation}$$

We used this model for smoking consumption only, due to the lack of suitable instruments for alcohol consumption. In fact, in order to separately identify the decision regarding participation (in our case: to smoke or not to smoke) from the level decision (how much to smoke) we need to find a variable or some variables (known as exclusion restrictions) which affect the decision of whether or not to smoke but do not affect the decision of how much to smoke. We chose to use a variable describing whether the cohort member's mother smoked during pregnancy or not. This variable is significantly correlated with individuals' selection into the smoking category[11], but is exogenous with respect of the amount of consumption.

[10] The sample selection model was first introduced by Heckman (1976, 1979) in the context of labour maket participation and wages of women.
[11] Much physiological and medical research in fact finds that adolescents are significantly more likely to smoke if their parents smoke (see for instance Ary et al. 2003, Harakeh et al. 2004, Hill et al. 2005, Jackson et al. 1997).

3.2 Longitudinal dynamics

As stressed above the link between education, skills and health-related outcomes may simply reflect individuals' unobserved characteristics that affect both educational attainment and health. Even if we can control for a large number of individuals' observable characteristics, there may still be unobserved factors that might bias the cross-section estimates. We are able to control for such unobserved heterogeneity by exploiting the longitudinal nature of our data and using a fixed effect estimator. In this way, all the time-invariant individual specific characteristics are taken into account through the inclusion of individual dummies. Moreover, the longitudinal dimension of the data is useful since it allows the possibility of determining the impact of previous event history on current behaviour. We therefore use the information on health-risky behaviours (smoking, drinking, and being obese and overweight) contained in the 1986 sweep of the BCS70 to control for initial condition (see Dorsett (1998), on the importance of accounting for the initial conditions in the econometric modeling of smoking). Since in 1986 cohort members were 16, we can reasonably think that this initial observation coincides with the starting time of health-risky behaviours.

Among the few economic papers focusing on this issue, Loureiro et al. (2006) show that the smoking transmission link is especially strong between parents and children of the same sex.

4. Data and descriptive statistics

Our empirical analysis relies on different sweeps of the British Cohort Study 1970.

The first part of our empirical analysis uses data from the 2004 sweep which includes 9665 individuals aged 34. The number of observations included in the regressions varies according to the availability of information for each variable (see Table 3).

As mentioned in the previous section, we will focus on three health-related outcomes: drinking behaviour, smoking status and Body Mass Index (BMI). Drinking behaviour is measured using daily (or weekly) alcohol consumption. At both 16 and 34, questions were asked about the number of units of alcohol that each participant had drunk over the past week. In order to compute daily alcohol consumption at 16, the total number of units drunk over the prior week was divided by the number of days on which they had been drinking (frequency). At age 34, daily alcohol consumption represents an estimate rather than an accurate reflection of daily alcohol consumption, as individuals were asked about the frequency of drinking rather than the exact number of days they had drunk on over the past week (see Appendix A).

Government guidelines defines 'binge drinking' as excessive alcohol consumption over a short period of time, normally considered a few hours or one night, which often leads to drunkenness (Parliamentary Office of Science and Technology 2005). Following established guidelines, binge drinking was computed for the sample at all age groups (i.e. 16 and 34). Binge drinking was defined as six or more units of alcohol in one day for females and eight or more units of alcohol in one day for males. We also considered another variable measuring relatively high alcohol use: we classified as 'heavy drinkers' the individuals whose weekly alcohol consumption is above the 90^{th} percentile of the overall distribution. While binge drinking refers to excessive drinking on one occasion only, heavy drinking is measuring cumulative drinking patterns across the week.

Our second dependent variable in terms of health behaviour regards smoking patterns. During both age sweeps, each respondent was asked whether he/she engaged in cigarette smoking and how many cigarettes he/she smoked on a daily basis. Smoking status was computed by assigning those who answered yes to cigarette smoking into the smoker category and those who answered no, whether or not they had been a smoker in the past, into the non-smoker category. We will also look at daily cigarette consumption which was directly drawn from the self-reported number of cigarettes smoked per day.

Our third and forth dependent variables indicate whether the respondent is overweight or obese. We therefore computed the BMI for each individual. This index is often used to determine how healthy an individual's weight is. For this study, BMI was calculated using the standard formula; that is, weight in kilograms divided by height in metres squared. Following clinical guidelines, we defined as overweight the individuals whose BMI was over 25 kg/m^2 and below 30 kg/m^2, while we defined as obese individuals with a BMI above 30 kg/m^2.

The table in Appendix B describes the distribution of the three outcomes variables in our sample.

As regards to our main variables of interest, we will consider the impact of both formal education and basic skills in literacy and numeracy. In 2004 all cohort members were assessed in terms of their basic skills using two tests for literacy and numeracy. The items in the tests are set at five levels of difficulty: Entry level 1, Entry level 2, Entry level 3, Level 1 and Level 2, the most difficult. The literacy was composed of 20 questions taken from the *Skills for Life Survey* (DfES 2003). Ten initial questions were introduced to screen individuals: when individuals scored lower that 6, they were asked 10 easier Entry Level 2 questions, while those who scored between 6 and 10 were given harder questions (five Level 1 and then five Level 2). The numeracy test was composed of 17 multiple-choice questions, asked to all individuals: five at Entry level 2, four at Entry level 3, five at Level 1 and three at Level 2 (for a detailed explanation of the tests' design, see Parsons and Bynner 2005).

Table 1 reports the distribution of cohort members' literacy and numeracy levels in 2004. It is worth noting that more than 8 per cent of individuals in our sample face severe literacy problems as their literacy skills are below the minimum target set by the Government (Level 1). In terms of numeracy, about 15 per cent of people have skills below the minimum target of Entry level 3. This picture is consistent with previous evidence documenting the poor record of adult UK population in terms of basic skills (see Moser Report 1999; the 2003 *Skill for Life Survey* and the 2006 Leitch Report).

Table 1: Distribution of cohort members' literacy and numeracy

	Literacy (% of sample)	Numeracy (% of sample)
Below Entry level 2	2.1	6.1
Entry level 2	2.2	8.9
Entry level 3 (*minimum target for numeracy*)	3.9	24.8
Level 1 (*minimum target for literacy*)	30.2	33.9
Level 2	61.7	26.3

The 2004 literacy and numeracy assessments also provide continuous measure of basic skills, based on the raw scores obtained. Drawing on these scores, we created a synthetic measure of basic skills by using Principal Component Analysis (PCA).

This measure of 'basic skills' is composed of two factors: the first factor contains eight lower tier literacy questions, four numeracy questions and one literacy screening question while the second factor is comprised of the remaining two lower tier literacy items, the remaining nine screening literacy questions, and nine upper tier literacy screening questions. The reliability of the basic skills factor is 0.88, indicating that it hangs together well (above 0.75 is considered good reliability)[12]. Subsequent correlation analysis (see Appendix C) suggests that while there is an overlap between literacy/numeracy and the basic skills factor,

[12] See Appendix C for a detailed explanation of the derivation of this measure and of the correlation analysis.

basic skills may be capturing something else which is hidden by the construction of two separate scales. This is echoed by the strong but not perfect correlation between basic skills and literacy/numeracy. Thus, it can be argued that combining the two factors into one measure of basic skills allows us to:

a) capture the shared variance among literacy and numeracy so that we can examine basic skills; and
b) provides us with a cleaner measure of basic skills than simply looking at the highest loading factor (i.e. the first factor).

In terms of formal education, the distribution of qualification levels in our sample is described in Table 2.

Table 2: Distribution of qualification levels

Qualification levels	Frequency	%	Cumulative
No qualification	899	9.34	9.34
Level 1 (e.g. CSE, low GCSEs, etc.)	1,457	15.14	24.48
Level 2 (e.g. good GCSEs, NVQ2, etc.)	3,173	32.97	57.45
Level 3 (e.g. A-levels, etc.)	884	9.19	66.64
Level 4 (e.g. degree, etc.)	2,605	27.07	93.7
Level 5 (e.g. MSc, PhD, etc.)	606	6.3	100
Total	9,624	100	

About 57 per cent of people have qualification levels equal or below the GCSE level. Almost 10 per cent of individuals have reached the A-level, and about 27 per cent obtained a degree or an equivalent level. Only 6 per cent of our sample has obtained a postgraduate degree. Again, this picture reflects the distribution of education at the national level (see Office for National Statistics 2005), confirming the good representativeness of our sample.

It is interesting to look at the relationship between formal education and basic skills assessments. While obviously basic skills increase as the qualification level gets higher, there is still a great variation of basic skills within each educational group. Figure 1 plots the average value of our basic skills indicator for each qualification level. We can note that higher qualification levels are associated with higher values of basic skills. The average basic skill measure for individuals with no qualifications is about 24, while the value for individuals with a postgraduate degree is about 31. Interestingly, the marginal increase in basic skills associated with higher qualification levels is decreasing: there are greater variations in basic skills among people with lower levels of qualification than among more educated individuals. This shows that the basic skill tests are designed to discriminate especially among low educated people. Figure 2 shows the distribution of basic skills by qualification levels within each educational group and underlines that there is significant variation in basic skills within each educational group. This suggests that education and basic skills, although related and sometime overlapping concepts, are in fact capturing different aspects of a person's human capital. In this sense, it is important to insert both the measures in our model.

Figure 1: Basic skills and qualification levels

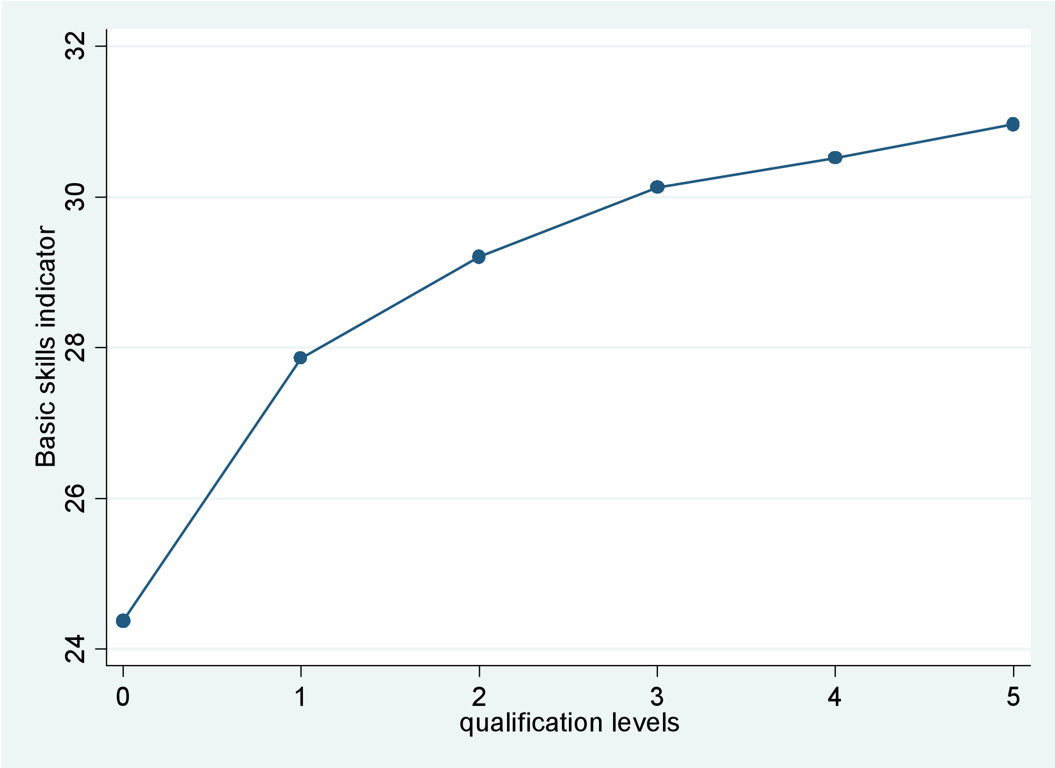
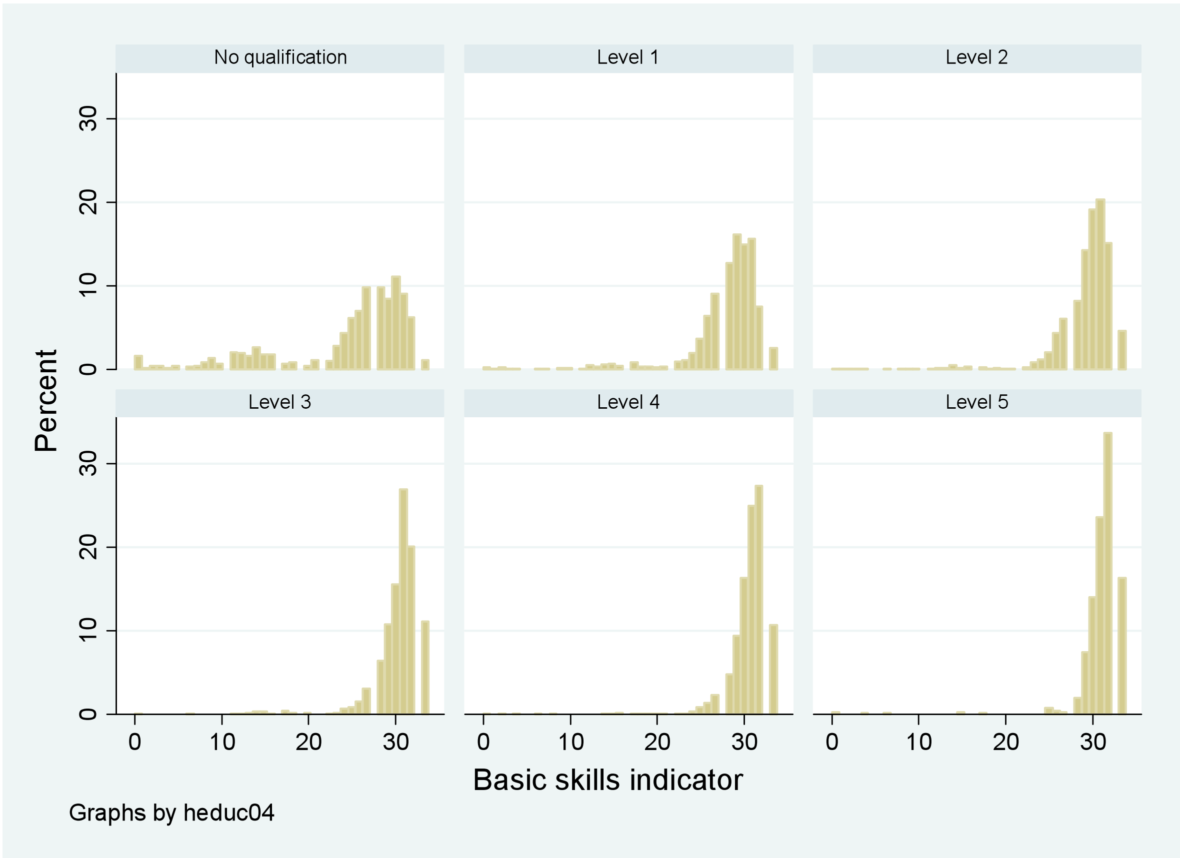


Figure 2: Distribution of basic skills by qualification levels



The next two graphs show the average values of our different health-related outcomes across different educational groups (Figure 3) and different basic skills quantiles (Figure 4). The two figures underline the reduction in detrimental health behaviours associated with skills improvements. Figure 3 shows that among people with no qualification, more than 50 per cent smoke, more than 40 per cent are heavy drinkers, about 30 per cent are binge drinkers and about 25 per cent are obese. These percentages reduce drastically as we look at people with more qualification: the percentage of smokers and heavy drinkers amongst the graduates is around only 15, that of binge drinker is around 20 and that of people classified as obese is less than 10. The same negative relationship is observable between the same health outcomes and different quintiles of basic skills. Individuals in the lower quintiles of the basic skill distribution tend to have worse health outcomes with respect to individuals in the upper part of the basic skill distribution. Among the different outcomes we are considering that of binge drinking seems to be the least related to education and basic skills. Overall, these figure point out that there is in fact a link between education and health-related behaviours. Whether this relationship depends on other characteristics that differently affect people with different levels of education or whether they imply a causal link will be the matter of the next section.

Before showing the results, we present Table 3 which summarises all the variables used in the analysis and provides some descriptive statistics.

Figure 3: Average health outcomes in different educational groups

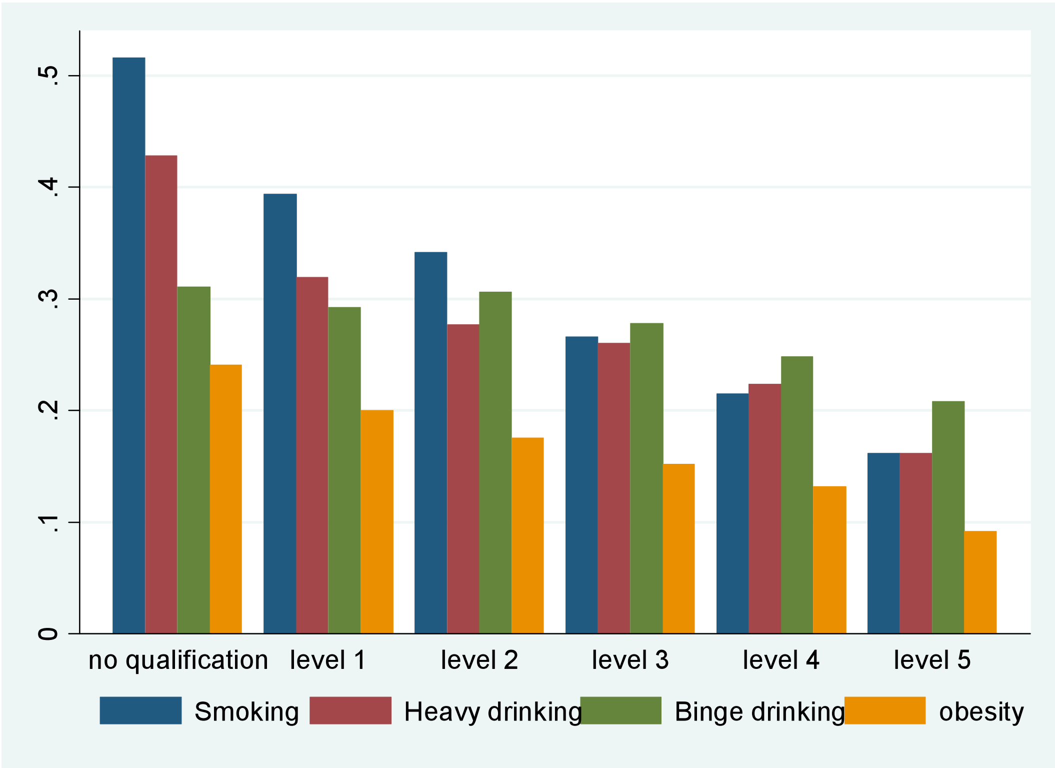


Figure 4: Average health outcomes in different basic skill quantiles

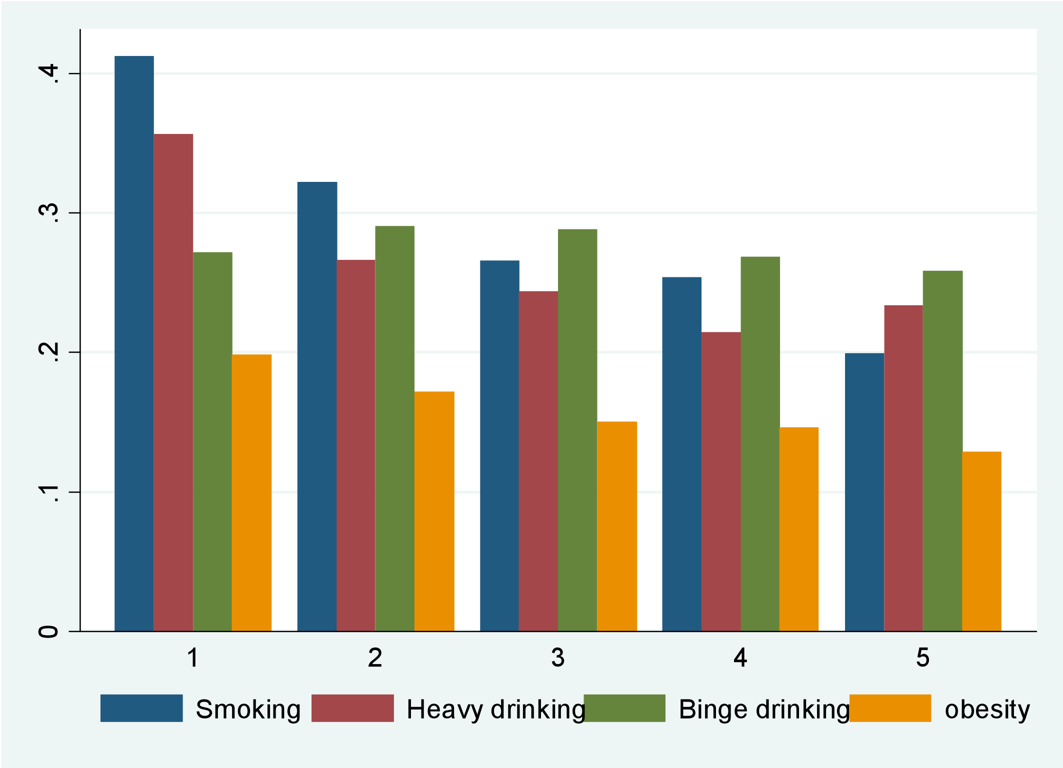


Table 3: Variables used in the analysis

Variable description	N*	Mean	Min.	Max.	N	Mean	Min.	Max.
	2004				1986			
Whether drinking	9588	0.19	0	1	4548	0.61	0	1
Daily alcohol consumption	9588	5.17 (5.26)	0	32	4548	2.31 (2.88)	0	16
Whether binge drinking	7745	0.28 (0.45)	0	1	4570	0.10 (0.29)	0	1
Heavy drinking	9665	0.28 (0.01)	0	1	11615	0.10 (0.01)	0	1
Whether smokes	9633	0.31 (0.46)	0	1	6129	0.26 (0.44)	0	1
Number of cigarettes a day	2366	14.84 (7.41)	1	80	899	8.91 (6.54)	1	45
Body mass index	9355	25.89 (4.93)	9.58	89.17	5723	21.26 (3.25)	10.80	67.58
Whether obese (BMI >30)	9355	0.17 (0.37)	0	1	5723	0.02 (0.13)	0	1
Whether overweight (25 < BMI <30)	9355	0.34 (0.47)	0	1	5723	0.08 (0.28)	0	1

Table 3: Variables used in the analysis (continued)

Variable description	N*	Mean	Min.	Max.	N	Mean	Min.	Max.
	2004				1986			
Measure of basic skills	9448	0 (1.29)	-6.27	1.75	3259	0 (1.54)	-8.17	2.79
Highest education level	9624	2.48 (1.42)	0	5	3492	1.42 (0.49)	0	2
Female	7745	0.53 (0.50)	0	1	6129	0.57 (0.49)	0	1
Whether CM** lives with a partner	9640	0.75 (0.44)	0	1	-	-	-	-
Full-time employment	9636	0.66 (0.47)	0	1	-	-	-	-
Part-time employment	9636	0.17 (0.38)	0	1	-	-	-	-
Frequency reading newspaper	9635	4.69 (1.32)	1	6	-	-	-	-
Life satisfaction scale	9594	7.40 (1.80)	0	10	-	-	-	-
Log weekly income	6755	2.01 (0.49)	-0.56	3.97	-	-	-	-
Locus of control	9561	0.30 (0.67)	0	3	-	-	-	-
Whether availability of magazines at home	-	-	-	-	6208	0.52 (0.50)	0	1

Notes: * N = numbers of observations. ** CM = cohort member.

5. Results

This section presents and discusses the estimates' outcomes. In particular, Section 5.1 focuses on the results of probit regressions, where we investigate the role of human capital on the probability of being a binge drinker, a smoker, obese and overweight. Section 5.2 then shows the results of Tobit and Heckman selection models to estimate the determinants of alcohol and cigarette consumption patterns. Finally, in Section 5.3, we show the results based on a longitudinal analysis.

5.1 Human capital and the occurrence of health risk-taking behaviours

The first tables show the results of the probit regressions on binge drinking and smoking (Table 4) and on BMI (Table 5). All the tables include the whole set of control variables. Each column differs according to the human capital variable we have inserted: first the education dummies and the constructed synthetic index of basic skills separately (columns 1, 4, 7 and 2, 5, 8 respectively) and then the two variables together in columns 3, 6 and 9. The tables report the value of the *marginal effect* of each variable on the probability of being a binge drinker, a smoker, obese or overweight. This allows a straightforward interpretation of the values in the tables as percentage point changes in these probabilities.

The results in Table 4 suggest that basic skills and education significantly affect the probability of being a heavy/binge drinker and a smoker. In particular, educational attainment affects both alcohol consumption and smoking status; those with higher levels of qualifications are significantly less likely to engage in heavy alcohol use and are less likely to smoke compared to individuals possessing no qualifications. Individuals with qualifications at Level 5 are almost 13 per cent (12 per cent) less likely to be binge drinkers (heavy drinkers) in comparison to individuals with no qualifications. The effect is even higher for smoking: the probability of being a smoker is 20 per cent lower for individuals with the highest qualification in comparison to those with no qualification at all. Also, in the case of smoking, having obtained a CSE (qualification Level 1) significantly reduces the likelihood of being a smoker.

As far as basic skills are concerned, our synthetic index of literacy and numeracy abilities is negatively related to smoking behaviour and heavy drinking as expected. The effect on smoking is only visible when the basic skills index is inserted alone (Table 4, column 5); this suggests that basic skills do not have an additional impact once the educational levels are taken into account. However, the inclusion of the basic skills indicator in the regression is still important as it controls for the part of ability that is not captured by education. The magnitude of the education dummies in fact decreases as we move from column 4 to column 6, meaning that if we do not control for literacy and numeracy abilities, the education variable may also pick up the impact of basic skills. Interestingly, basic skills are significantly negatively related to heavy drinking behaviour even when education is controlled for (see column 9). This means that having better basic skills decreases the odds of being a heavy drinker, conditional on the level of education.

Similar results apply on the probability of being obese (see Table 4) both formal education and basic skills seem to reduce significantly the odds of being obese when they are inserted separately. However, once the two variables are jointly inserted, the effect of education is dominant. We also found that neither basic skills nor education have any significant impact on the probability of being overweight. We interpret this result by considering that being overweight is not as harmful for health and therefore, the link with human capital is weaker than between obesity and human capital.

These results are robust to the inclusion of a vast array of potential confounding factors, supporting that the impact of education and basic skills is not conditional on personal characteristics – such as sex; whether living with a partner or not; socioeconomic background – type of occupations, whether full-time employed, log income, availability of newspapers and magazines at home; as well as measures of life satisfaction and locus of control. While we inserted these variables only as controls – as they may mediate the effect of education – it is still interesting to comment on their impact on health behaviour, our variables of interest.

It seems that females are significantly less likely to have problems with alcohol and are less likely to be overweight. No significant differences exist between men and women with respect to the propensity of smoking and being obese.

Full-time employment tends to increase the probability of drinking, smoking and being obese. A tentative explanation for these results could be that full-time workers are more likely to be stressed by their jobs, turning to alcohol, cigarettes and food as coping strategies. Additionally, some corporate cultures encourage interacting in a social environment, perhaps leading to increased drinking and food consumption.

The different occupation dummies are positive and significant in the regressions on binge drinking and on smoking: this suggests that doing jobs different from 'high managerial and professional' (the most qualified one) increases the probability of risky health behaviours. However, job type is only weakly related to an individual's BMI. Smoking behaviour and obesity are also affected by net income which seems to reduce the likelihood of being a smoker and obese. All these variables may be considered as proxies for social class and thus overall, the results depict a positive association between social class and health outcomes.

We also find that living with a partner decreases the probability of binge drinking and of being a smoker, but tends to be associated with a higher probability of being obese and overweight. It is possible that cohabiting with a partner who does not drink excessively or smoke is protective. The number of children seems to significantly reduce the probability of being binge drinker. It is also positively associated with the probability of being obese.

Our results also suggest that the (self-assessed) quality of the local health service does matter in terms of individuals' health outcomes (in particular, it seems to reduce the probability of smoking and binge drinking). The reason could be that well-functioning health services help to increase people's awareness on the potential harm from health-risky behaviours and thus lead individuals to pay more attention to their lifestyle.

As far as the psychological variables are concerned, it seems that locus of control and life satisfaction significantly affects the probability of smoking (those with a high *external* locus of control and with lower life satisfaction are more likely to smoke).

However, it is important to underline that while we can give an interpretation to the basic skills coefficient, we are more cautious in interpreting the other control variables because of issues of reverse causality.

Table 4: Probit regressions on smoking, binge drinking and heavy drinking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Smoking			***Binge drinking***			***Heavy drinking***		
Basic skills		-0.0232***	-0.00691		0.0113	0.0209**		-0.0258***	-0.0201***
		(0.0077)	(0.0081)		(0.0083)	(0.0088)		(0.0072)	(0.0076)
Educ-Level 1	-0.0640***		-0.0570**	-0.0347		-0.0472*	-0.0440*		-0.0332
	(0.025)		(0.026)	(0.027)		(0.028)	(0.024)		(0.026)
Educ-Level 2	-0.113***		-0.105***	-0.0358		-0.0522*	-0.0625***		-0.0476**
	(0.023)		(0.024)	(0.026)		(0.027)	(0.023)		(0.024)
Educ-Level 3	-0.136***		-0.129***	-0.0610**		-0.0784***	-0.0623**		-0.0466*
	(0.023)		(0.024)	(0.029)		(0.029)	(0.026)		(0.028)
Educ-Level 4	-0.185***		-0.178***	-0.0890***		-0.107***	-0.0700***		-0.0549**
	(0.023)		(0.024)	(0.026)		(0.027)	(0.024)		(0.026)
Educ-Level 5	-0.202***		-0.197***	-0.133***		-0.148***	-0.123***		-0.110***
	(0.020)		(0.021)	(0.027)		(0.027)	(0.024)		(0.026)
Female	0.0234	0.0105	0.0238	-0.0947***	-0.101***	-0.0935***	-0.0242*	-0.0286**	-0.0253*
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
Living with partner	-0.0878***	-0.0846***	-0.0879***	-0.0874***	-0.0859***	-0.0878***	-0.00942	-0.00758	-0.00819
	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)
Full-time employment	0.0285	0.0275	0.0296	0.0474**	0.0454**	0.0471**	-0.0400**	-0.0395**	-0.0394**
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Lower managerial	-0.00189	0.0189	-0.00249	0.0241	0.0413**	0.0263	0.00887	0.0127	0.00677
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)
Intermediate oc.	-0.0164	0.0297	-0.0180	0.0401	0.0768***	0.0434*	0.00302	0.0133	-0.00107
	(0.024)	(0.025)	(0.024)	(0.026)	(0.025)	(0.026)	(0.023)	(0.023)	(0.023)
Technical oc.	0.0843***	0.152***	0.0800***	0.0545**	0.103***	0.0591**	0.0506*	0.0624**	0.0430*
	(0.028)	(0.027)	(0.028)	(0.027)	(0.027)	(0.027)	(0.026)	(0.025)	(0.026)
Semi-routine oc.	0.0488*	0.116***	0.0450	0.0406	0.0951***	0.0510*	0.0298	0.0400	0.0202
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.029)	(0.026)	(0.026)	(0.026)
Routine oc.	0.145***	0.234***	0.141***	0.0564*	0.122***	0.0684**	0.0433	0.0559*	0.0325
	(0.034)	(0.033)	(0.034)	(0.032)	(0.032)	(0.033)	(0.030)	(0.029)	(0.030)
Life satisfaction	-0.00608	-0.00739*	-0.00677	0.00227	0.00169	0.00209	-0.00925**	-0.0103**	-0.00976**
	(0.0043)	(0.0043)	(0.0043)	(0.0045)	(0.0045)	(0.0045)	(0.0041)	(0.0041)	(0.0041)
Net Income	-0.0338**	-0.0488***	-0.0326**	0.0112	-0.00124	0.00933	-0.0304**	-0.0330**	-0.0297**
	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
Locus of control	0.0383***	0.0383***	0.0364***	-0.0133	-0.0130	-0.0138	0.000347	-0.000658	-0.000994
	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
Frequency of reading newspaper	0.00926*	0.0102**	0.00943*	0.0144***	0.0153***	0.0146***	-0.00775*	-0.00770*	-0.00798*
	(0.0049)	(0.0049)	(0.0049)	(0.0050)	(0.0050)	(0.0050)	(0.0046)	(0.0046)	(0.0046)
Quality of local health services	-0.0178***	-0.0194***	-0.0178***	-0.00992	-0.0109*	-0.00986	-0.0161***	-0.0161***	-0.0160***
	(0.0062)	(0.0062)	(0.0062)	(0.0063)	(0.0063)	(0.0063)	(0.0059)	(0.0059)	(0.0059)
Number of children	-0.00840	-0.00449	-0.00756	-0.0183**	-0.0165**	-0.0181**	-0.00585	-0.00527	-0.00634
	(0.0071)	(0.0071)	(0.0071)	(0.0072)	(0.0072)	(0.0072)	(0.0066)	(0.0066)	(0.0067)
Observations	5433	5419	5414	5433	5419	5414	5433	5419	5414

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 5: Probit regressions on being obese and overweight

	(1)	(2)	(3)	(4)	(5)	(6)
	Obesity			***Overweight***		
Basic skills		-0.0131**	-0.00780		-0.00358	-0.00190
		(0.0063)	(0.0067)		(0.0087)	(0.0091)
Educ-Level 1	-0.0308		-0.0255	0.0341		0.0353
	(0.020)		(0.021)	(0.032)		(0.033)
Educ-Level 2	-0.0451**		-0.0383*	0.0163		0.0175
	(0.019)		(0.020)	(0.029)		(0.030)
Educ-Level 3	-0.0315		-0.0246	-0.0153		-0.0132
	(0.022)		(0.024)	(0.034)		(0.035)
Educ-Level 4	-0.0714***		-0.0652***	-0.0146		-0.0127
	(0.020)		(0.021)	(0.030)		(0.032)
Educ-Level 5	-0.0973***		-0.0926***	-0.0573		-0.0548
	(0.018)		(0.019)	(0.035)		(0.037)
Female	-0.0166	-0.0229*	-0.0168	-0.191***	-0.194***	-0.192***
	(0.012)	(0.012)	(0.012)	(0.016)	(0.016)	(0.016)
Living with partner	0.00766	0.00815	0.00708	0.0426**	0.0448***	0.0440***
	(0.013)	(0.013)	(0.013)	(0.017)	(0.017)	(0.017)
Full-time employment	0.0606***	0.0564***	0.0602***	-0.0116	-0.0120	-0.0129
	(0.014)	(0.014)	(0.014)	(0.021)	(0.021)	(0.021)
Lower managerial occupations	0.00558	0.0138	0.00503	0.0402**	0.0473**	0.0400**
	(0.016)	(0.016)	(0.016)	(0.020)	(0.020)	(0.020)
Intermediate occupations	0.0187	0.0377*	0.0162	-0.000486	0.0194	-0.0000742
	(0.021)	(0.021)	(0.021)	(0.026)	(0.026)	(0.026)
Technical occupations	0.00757	0.0310	0.00459	0.0476*	0.0715***	0.0473*
	(0.021)	(0.022)	(0.021)	(0.028)	(0.027)	(0.028)
Semi-routine occupations	0.0317	0.0523**	0.0259	0.0381	0.0611**	0.0377
	(0.024)	(0.024)	(0.023)	(0.029)	(0.029)	(0.030)
Routine occupations	0.0132	0.0428	0.00792	0.00151	0.0291	0.00524
	(0.025)	(0.026)	(0.025)	(0.032)	(0.031)	(0.032)
Life satisfaction	-0.00576	-0.00584*	-0.00577	-0.00445	-0.00486	-0.00423
	(0.0035)	(0.0035)	(0.0035)	(0.0047)	(0.0047)	(0.0047)
Net income	-0.0518***	-0.0588***	-0.0517***	0.00875	0.00227	0.00771
	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)	(0.016)
Locus of control	-0.00521	-0.00342	-0.00479	-0.0154	-0.0161	-0.0168
	(0.0094)	(0.0095)	(0.0095)	(0.013)	(0.013)	(0.013)
Frequency of reading newspaper	0.00125	0.00159	0.000915	0.00339	0.00388	0.00346
	(0.0040)	(0.0040)	(0.0040)	(0.0052)	(0.0052)	(0.0052)
Quality of local health services	0.00652	0.00579	0.00635	0.0118*	0.0116*	0.0119*
		(0.0052)	(0.0052)	(0.0067)	(0.0067)	(0.0067)
Number of children	0.0185***	0.0188***	0.0185***	0.00668	0.00714	0.00575
	(0.0056)	(0.0056)	(0.0056)	(0.0075)	(0.0075)	(0.0075)
Observations	5292	5279	5274	5292	5279	5274

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Additional analyses to differentiate these results by gender and level of qualifications were performed, but the results were hardly affected and we do not report them here. We also investigated whether basic skills complement education by introducing interaction dummies. The combined effect of education and basic skills were non-significant, suggesting the impact of basic skills is not differentiated by education groups.

5.2 Human capital and the quantity of smoking and drinking

In this section, we explore the impact of basic skills on the *quantity* of drinking and smoking (not only on the probability of being a heavy drinker or smoker). In this way, we explore the entire variability in health outcomes (instead of reducing them to binary outcomes). Moreover, the role of human capital may differ in the decision of taking a health-risky behaviour, such as smoking or binge drinking, and in the consumption decision. It is thus worthwhile to investigate this role as well.

Table 6 reports the results for Tobit regressions on alcohol consumption. As explained above, the likelihood function for a Tobit model involves both the process that determines whether the outcome variable is fully observed or not and the process that determines the score on the dependent variable for individuals whose outcome is fully observed. The results suggest that the variables affecting the probability of being a binge drinker and a heavy drinker also affect in the same direction the amount of alcohol consumed. The more educated the individuals, the lower the amount of alcohol they drink. The incremental effect of basic skills is not significant in this case. Our results also suggest that being a female and living with a partner significantly reduces the amount of alcohol consumption whereas being employed in unskilled routine jobs is associated with higher weekly alcohol consumption.

Table 6: Tobit regressions on weekly alcohol consumption

	(1)	(2)	(3)
	Weekly alcohol units		
Basic skills		0.348	0.511*
		(0.29)	(0.30)
Educ-Level 1	-0.0436		-0.249
	(1.00)		(1.03)
Educ-Level 2	-0.0485		-0.335
	(0.92)		(0.96)
Educ-Level 3	-2.044*		-2.439**
	(1.09)		(1.13)
Educ-Level 4	-0.872		-1.294
	(0.98)		(1.03)
Educ-Level 5	-2.281*		-2.719**
	(1.17)		(1.21)
female	-7.686***	-7.785***	-7.682***
	(0.49)	(0.49)	(0.49)
Living with partner	-1.301**	-1.299**	-1.319**
	(0.52)	(0.53)	(0.53)
Full-time employment	1.217*	1.211*	1.181*
	(0.66)	(0.66)	(0.66)
Lower managerial	0.0596	0.331	0.104
occupations	(0.59)	(0.59)	(0.59)
Intermediate	-1.065	-0.539	-1.035
occupations	(0.80)	(0.77)	(0.80)
Technical	1.444*	2.273***	1.501*
occupations	(0.83)	(0.79)	(0.84)
Semi-routine	0.222	1.076	0.368
occupations	(0.88)	(0.85)	(0.89)
Routine	1.553	2.580***	1.774*
occupations	(0.99)	(0.95)	(1.00)
Life satisfaction	-0.369**	-0.382***	-0.383***
	(0.15)	(0.15)	(0.15)
Net income	1.877***	1.629***	1.825***
	(0.47)	(0.47)	(0.47)
Locus of control	-0.565	-0.586	-0.613
	(0.39)	(0.40)	(0.40)
Frequency of	0.140	0.141	0.135
reading newspaper	(0.17)	(0.17)	(0.17)
Quality of local health services	-0.711***	-0.747***	-0.723***
	(0.21)	(0.21)	(0.21)
Number of children	-1.103***	-1.044***	-1.089***
	(0.24)	(0.24)	(0.24)
Constant	19.03***	18.62***	19.55***
	(2.24)	(2.11)	(2.25)
Observations	4520	4507	4503

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

In Table 7, we use the same methodology (Tobit regressions) to estimate the determinants of the average amount of cigarettes smoked weekly. Regressions are shown in columns 1, 4 and 7. This model allows us to investigate the consumption of cigarettes taking into account the fact that a large number of respondents do not smoke cigarettes at all (i.e.76 per cent are non-smokers). We observe in column 1 that basic skills are negatively correlated with cigarette consumption. The same applies to qualifications levels. Each level of qualifications is associated with strong and significant reductions in consumption and the size of the coefficients increases with the qualification levels. When introduced together with education levels, the residual effect of basic skills disappears (column 7).

The results based on Tobit estimations suggest alternative models of estimation, where both the decision to smoke and the quantity of cigarettes consumed are investigated. In line with previous attempts in literature on smoking (see for example Adda and Cornaglia 2006) we perform a joint estimation of the probability to smoke (yes or no) and the average amount of cigarettes smoked per day. We introduce the same variables as previously in both equations. For identification that does not rely upon the normality assumption, we need to introduce a variable correlated with the decision to smoke and not the consumption (see Vella 1998). As an exclusion restriction, we use whether the mother of the cohort member was smoking during her pregnancy, as the smoking pattern of parents is often found to be strongly related with their children's decision to smoke (Loureiro et al. 2006). We are only moderately successful in this attempt as the coefficient for the non-random inclusion in the smoking consumption equation (i.e. the inverse Mills ratio) is significant only in one out of three equations. The coefficient on whether the mother smoked during pregnancy is, however, highly significant for the decision to smoke (P-value of 0 until up to three digits after comma).

More interestingly for our purpose is that human capital appears to have a significant (negative) impact on smoking initiation, but less impact on the amount of cigarettes smoked. When basic skills are introduced alone (column 2), they appear strongly significant and negatively related to the decision to smoke and non-significant with cigarette consumption. When qualification levels are introduced, it appears that it is towards the bottom of the qualification distribution that increased human capital explains the decision (not) to smoke but not the amount smoked. Higher up the qualifications level, education decreases both the decision to smoke and the amount smoked.

The effect of accumulated human capital is measured by education only and the basic skills effect is not significant anymore in column 8, underlining that for given levels of qualification basic skills do not have an independent effect.

It is also interesting to observe that the socio-economic position (measured by occupations of decreasing level of responsibility and independence in execution) is significantly related to the decision to smoke or not (people in lower occupations are more likely to smoke), but unrelated to cigarette consumption.

Also, individuals with lower levels of internal control tend to choose more often to smoke, but they tend to smoke on average fewer cigarettes than those with higher levels of internal control.

Table 7: Tobit and Heckman selection models on cigarette consumption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Tobit	**Heckman's selection model**		**Tobit**	**Heckman's selection model**		**Tobit**	**Heckman's selection model**	
	Cig per day	*Decision to smoke*	*Cig per day*	*Cig per day*	*Decision to smoke*	*Cig per day*	*Cig per day*	*Decision to smoke*	*Cig per day*
Basic skills	-1.641***	-0.086***	-0.234				-0.594	-0.0380	-0.594
	(0.44)	(0.025)	(0.23)				(0.46)	(0.027)	(0.46)
Educ-Level 1				-4.106***	-0.214**	-0.980	-3.617**	-0.173*	-3.617**
				(1.47)	(0.090)	(0.73)	(1.53)	(0.093)	(1.53)
Educ-Level 2				-7.133***	-0.366***	-1.179	-6.538***	-0.317***	-6.538***
				(1.36)	(0.083)	(0.72)	(1.44)	(0.087)	(1.44)
Educ-Level 3				-9.669***	-0.476***	-1.812*	-9.039***	-0.425***	-9.039***
				(1.75)	(0.10)	(0.99)	(1.82)	(0.11)	(1.82)
Educ-Level 4				-13.73***	-0.695***	-2.076**	-13.19***	-0.648***	-13.19***
				(1.56)	(0.093)	(1.01)	(1.65)	(0.098)	(1.65)
Educ-Level 5				-18.38***	-0.901***	-2.791*	-17.77***	-0.848***	-17.77***
				(2.23)	(0.13)	(1.57)	(2.29)	(0.13)	(2.29)
Female	-0.0727	0.0531	-2.437***	0.775	0.102**	-2.289***	0.786	0.102**	0.786
	(0.89)	(0.050)	(0.48)	(0.88)	(0.051)	(0.49)	(0.89)	(0.051)	(0.89)
Living with partner	-4.708***	-0.230***	-0.144	-4.787***	-0.244***	-0.215	-4.827***	-0.246***	-4.827***
	(0.93)	(0.053)	(0.51)	(0.92)	(0.053)	(0.51)	(0.92)	(0.053)	(0.92)
Full-time employment	2.946**	0.169***	0.280	2.910**	0.170***	0.388	2.966***	0.175***	2.966***
	(1.15)	(0.065)	(0.65)	(1.14)	(0.066)	(0.65)	(1.14)	(0.066)	(1.14)
Lower managerial occupations	3.181***	0.113*	0.635	1.418	0.0339	0.339	1.365	0.0298	1.365
	(1.23)	(0.068)	(0.76)	(1.23)	(0.069)	(0.76)	(1.23)	(0.070)	(1.23)
Intermediate occupations	3.219**	0.151*	-0.369	-0.334	-0.0160	-0.642	-0.489	-0.0222	-0.489
	(1.50)	(0.083)	(0.89)	(1.53)	(0.086)	(0.89)	(1.54)	(0.087)	(1.54)
Technical occupations	10.74***	0.478***	1.441	5.929***	0.260***	0.972	5.648***	0.245***	5.648***
	(1.49)	(0.083)	(0.94)	(1.54)	(0.088)	(0.92)	(1.55)	(0.088)	(1.55)
Semi-routine occupations	9.768***	0.471***	1.048	4.988***	0.254***	0.603	4.664***	0.238***	4.664***
	(1.55)	(0.086)	(0.96)	(1.59)	(0.090)	(0.94)	(1.60)	(0.091)	(1.60)
Routine occupations	15.84***	0.743***	2.203**	10.02***	0.485***	1.598	9.739***	0.467***	9.739***
	(1.67)	(0.094)	(1.07)	(1.71)	(0.099)	(1.02)	(1.73)	(0.10)	(1.73)
Frequency of reading newspaper	0.497*	0.0262	-0.0967	0.430	0.0226	-0.0438	0.435	0.0239	0.435
	(0.29)	(0.016)	(0.15)	(0.28)	(0.016)	(0.15)	(0.29)	(0.016)	(0.29)
Life satisfaction	-0.634**	-0.0342**	-0.147	-0.528**	-0.0287**	-0.163	-0.563**	-0.0313**	-0.563**
	(0.25)	(0.014)	(0.14)	(0.25)	(0.014)	(0.13)	(0.25)	(0.014)	(0.25)
Net income	-3.754***	-0.185***	-0.249	-2.618***	-0.134***	-0.0439	-2.545***	-0.128***	-2.545***
	(0.86)	(0.048)	(0.50)	(0.86)	(0.049)	(0.49)	(0.87)	(0.049)	(0.87)
Locus of control	1.938***	0.126***	-0.632**	1.986***	0.136***	-0.605*	1.887***	0.126***	1.887***
	(0.64)	(0.037)	(0.32)	(0.62)	(0.037)	(0.32)	(0.63)	(0.037)	(0.63)
Quality of local health services	-1.014***	-0.0488**	-0.335	-0.897**	-0.0425**	-0.315	-0.902**	-0.0434**	-0.902**
	(0.37)	(0.021)	(0.20)	(0.37)	(0.021)	(0.20)	(0.37)	(0.021)	(0.37)
Number of children	0.417	0.00962	0.253	0.152	-0.00463	0.240	0.197	-0.00061	0.197
	(0.42)	(0.024)	(0.23)	(0.41)	(0.024)	(0.22)	(0.42)	(0.024)	(0.42)
Mother smoking during pregnancy		0.252***			0.218***			0.215***	
		(0.044)			(0.044)			(0.044)	
Constant	-4.724	-0.396*	20.13***	3.872	0.0323	20.51***	3.591	0.00259	3.591
	(3.79)	(0.21)	(2.34)	(3.86)	(0.22)	(2.25)	(3.89)	(0.22)	(3.89)
Lambda			-1.783			-1.388			-1.513
			(1.022)			(1.082)			(1.072)
Log likelihood	-6788.21	-5981.664		-6766.14	-5965.42		-6730.73	-5933.08	
Observations	5064	4725		5077	4736		5059	4720	

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

5.3 Human capital and changes in health-risky behaviours over the lifetime

As explained above (see Section 3.2) we exploit the longitudinal nature of our data by estimating an equation which includes a lagged dependent variable (at age 16 in 1986) and by using a linear probability model with a Fixed Effect (FE) estimator. The estimates' outcomes are summarised in Table 8 which reports the result for each health-related outcome using the two methods. Columns 1 and 2 refer to the probability of being a binge drinker, columns 3 and 4 to the probability of being a heavy drinker, columns 5 and 6 of being a smoker, 7 and 8 of being obese and 9 and 10 of being overweight. Columns 1, 3, 5, 7 and 9 report the results for the probit model including a baseline variable (the dependent variable measured at age 16) and with the full set of control variables (those coefficients are not reported in the table). The results confirm the findings of the previous tables, highlighting the strong and significant role of education. In this case, we can interpret the coefficients as the marginal effects on the probability of changing behaviour between 1986 and 2004. It seems that having high educational levels significantly decreases the probability of engaging in health-risky behaviours at 34, given past behaviour at 16 (this is not the case, however, for changes in the probability of being overweight which resulted to be non-significantly related to education). As in the cross-section regressions, we find that basic skills do not have any significant additional impact on health once education is accounted for (with the exception of heavy drinking, when they enter the regression with a significant negative sign). This means that individuals with higher literacy and numeracy are less likely to be heavy drinkers at age 34, even after controlling for their past behaviour. Interestingly, it appears that the baseline measure is significant in predicting behavioural patterns at 34 only in case of obesity and heavy drinking. The probability of being obese and a heavy drinker is strongly and positively influenced by the obesity status and drinking behaviour at age 16. Instead, for the other three variables, behaviour at age 16 is not significantly related with adult outcomes.

Results of the fixed effect estimates are reported in columns 2, 4, 6, 8 and 10. Given that we are considering two periods only, the results can be interpreted as in the first differences estimation. In other words, the observed coefficients express the change in the probability of experiencing a specific health outcome as a result of changes in education and basic skills. In this way, we are netting out all the unobserved (time invariant) individuals' characteristics. The results confirm the important role of education in affecting health-related behaviours: the negative signs of the education dummies reveal that an increase in the education levels would reduce the probability of engaging in dangerous health behaviour. This effect seems to be particularly high and significant in the case of smoking.

Table 8: Longitudinal analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Binge drinking		*Heavy drinking*		*Smoke*		*Obese*		*Overweight*	
	Probit with baseline	FE	Probit with baseline	FE	Probit with baseline	FE	Probit with baseline	FE	Probit with baseline	FE
Binge [-1]	0.0450									
	(0.033)									
Heavy drink [-1]			0.0535***							
			(0.013)							
Smoke [-1]					0.0238					
					(0.018)					
Obese [-1]							0.455***			
							(0.068)			
Overweight [-1]									-0.00210	
									(0.033)	
Basic skills	0.0118	0.0418*	-0.0172**	0.0247	-0.0131	0.0195	-0.00946	0.00651	-0.0174	0.0242
	(0.013)	(0.024)	(0.0076)	(0.030)	(0.011)	(0.028)	(0.0093)	(0.019)	(0.013)	(0.028)
Educ-Level 1	-0.0437	-0.248*	-0.0289	0.0110	-0.0303	-0.363***	-0.0213	-0.266***	0.0498	-0.154
	(0.051)	(0.13)	(0.026)	(0.17)	(0.039)	(0.13)	(0.032)	(0.095)	(0.051)	(0.14)
Educ-Level 2	-0.0479	-0.128	-0.0391	-0.00964	-0.0639*	-0.342**	-0.0336	-0.103	0.0374	-0.116
	(0.050)	(0.13)	(0.025)	(0.17)	(0.037)	(0.14)	(0.031)	(0.096)	(0.046)	(0.14)
Educ-Level 3	-0.0787	-0.0156	-0.0376	-0.0536	-0.0820**	-0.395***	-0.0308	-0.0595	0.0632	0.115
	(0.050)	(0.14)	(0.029)	(0.18)	(0.037)	(0.15)	(0.034)	(0.10)	(0.055)	(0.15)
Educ-Level 4	-0.0888*	0.0356	-0.0440	-0.141	-0.140***	-0.437***	-0.0669**	-0.0518	0.0419	0.0930
	(0.052)	(0.14)	(0.027)	(0.17)	(0.037)	(0.14)	(0.032)	(0.097)	(0.049)	(0.14)
Educ-Level 5	-0.148***	-0.00199	-0.0989***	-0.0880	-0.153***	-0.443***	-0.0907***	-0.100	0.00316	0.0930
	(0.044)	(0.14)		(0.18)	(0.031)	(0.14)	(0.028)	(0.10)	(0.057)	(0.15)
Controls	✓		✓		✓		✓		✓	
Observations	2425	1632	5389	1632	3186	1714	2824	1194	2824	1194
Number of individuals		816		816		857		597		597

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
FE stands for individual fixed effect models and standard errors are in parentheses. Education is measured at age 16 as the highest level attained at this age, it takes therefore only 3 values: 0 = no qualifications at all; 1: = academic qualification at Level 1 attained; 2 = academic qualification at level 2 attained.

6. Conclusions

This report aimed at investigating how human capital human (measured by highest education levels and basic skills) is related to three health behaviour outcomes (drinking, smoking and weight). The dataset used in the analyses was the BCS70, a survey of over 18, 000 individuals born in 1970 who have been interviewed at various time points since birth: at age 5, 10, 16, 26, 30 and 34. The main focus of the analysis was on health-risky behaviours measured at age 34.

Our results suggest that education significantly reduced the probability of being a smoker (and smoking intensity), of being a binge drinker and of being classified as obese. These results hold after controlling for a large number of potential confounding variables, such as gender, occupational roles, work status, life satisfaction, locus of control and income. Using a longitudinal analysis over the age 16 to 34, we also find that past behaviours do have an important impact on the current one for the probability of being a heavy drinker and being obese.

Thus, the analysis provided further support that education is an important factor in the generation of healthier lifestyles in the UK. Our results indicate that basic skills are also important in the production of good health but we do not find any effect of basic skills once the highest education level has been introduced. This finding does not imply that basic skills have no impact on health, but possibly that our index of basic skills is not capturing health improving skills beyond the ones measured by education levels.

Overall, this report provides support for policies aimed at increasing adult skills in order to improve health-related outcomes in the UK. Our results appear to apply uniformly across the education distribution and do not differ significantly by gender.

References

Adams, S. (2002) 'Educational attainment and health: Evidence from a sample of older adults', *Education Economics*, vol. 10, pp.97–109.

Adda, J. and Cornaglia, F. (2006) 'Prices, Cigarette Consumption and Smoking Intensity', *American Economic Review*, vol. 96 (4), pp. 1013–1028.

Angrist, J. and Krueger, A. (1991) 'Does Compulsory School Attendance Affect Schooling and Earnings?', *Quarterly Journal of Economics*, 106(4), 979–1014.

Ary, D.V., Duncan, T.E., Duncan, S.C. and Hops, H. (2003) 'Adolescent Problem Behavior: the Influence of Parents and Peers', *Behavior Research and Therapy*, vol. 37, pp. 217–23.

Baker, D. (2006) 'The Meaning and the Measure of Health Literacy', *Journal of General Internal Medicine*, 21(8), 878–883.

Cutler, D.M. and Lleras-Muney, A. (2006) 'Education and Health: Evaluating Theories and Evidence', NBER Working Paper N. 12352, National Bureau of Economic Research.

Chandola, T., Clark, P., Morris, J. and Blane, D. (2006) 'Pathways between education and health: a causal modelling approach', *Journal of the Royal Statistical Association,* Series A, vol. 169 (2), 337–59.

de Coulon, A., Marcenaro-Gutierrez, O. and Vignoles, A. (2007) *The Value of Basic Skills in the British Labour Market*. London: Centre for Economic Performance (CEP), London School of Economics.

de Coulon, A. Meschi, E. and Vignoles, A. (2008) *Parents' basic skills and children's cognitive outcomes, CEE discussion paper, N 0104.* London: Centre for Economic Performance (CEP), London School of Economics.

DfEE (1999) *Improving literacy and numeracy: a fresh start*. Great Britain Working Group on Post-School Basic Skills chaired by Sir Claus Moser. London: Department for Education and Employment.

DfES (2003) *The Skills for Life Survey: A national needs and impact survey of literacy, numeracy and ICTskills*, DfES Research Report 490, available: www.dcsf.gov.uk/research/data/uploadfiles/rr490.pdf

DeWalt, D.A., Berman, N.D., Sheridan, S., Lohr, K.A. and Pignone, M.P. (2004) 'Literacy and health outcomes: A systematic review of the literature', *Journal of General Internal Medicine,* vol. 19, pp. 1228–1239.

DIUS (2009) Skills for Life: Changing lives. London: Department of Innovation, Universities and Skills.

Erbsland, M., Ried, M. and Ulrich, V. (1995): 'Health, health care, and the environment. Econometric evidence from German micro data', *Health Economics*, vol. 4, pp. 169–182.

Fuchs, V. (2004): 'Reflections on the socio-economic correlates of Health', *Journal of Health Economics*, Vol. 23, pp. 653–661.
Gilleskie, D.B. and Harrison, A.L. (1998): 'The effect of endogenous health inputs on the relationship between health and education', *Economics of Education Review*, vol. 17, pp. 279–297.

Groot, W. and Maassen van den Brink, H. (2007) 'The health effects of education', *Economics of Education Review*, vol. 26(2), pp. 186–200.

Grossman, M. (1972) 'On the concept of health capital and the demand for health', *Journal of Political Economy*, vol. 80, pp. 223–255.

Grossman, M. (2005) *Education and Non-market Outcomes*, NBER Working Papers, N. 11582, National Bureau of Economic Research.

Grossman, M. and Kaestner, R. (1997) 'Effects of education on health', in: J.R. Behrman and N. Stacey (eds) *The Social Benefits of Education.* Ann Arbor, MI: University of Michigan Press.

Harakeh, Z., Scholte, R., Vermulst, A., de Vries, H. and Engels, R. (2004) 'Parental Factors and Adolescents' Smoking Behavior: An Extension of the Theory of Planned Behavior', *Preventive Medicine,* 39(5), 951961.

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 Social Measurement*, vol. 5, pp. 475–492.

Heckman, J. (1979) 'Sample Selection Bias as a Specification Error", *Econometrica*, vol. 47, pp. 53–161.

Hill, K.G., Hawkins, J.D., Catalano, R.F., Abbott, R.D. and Guo, J. (2005) 'Family Influences on the Risk of Daily Smoking Initiation', *Journal of Adolescent Health,* vol. 37(3), pp. 202–210.

Jackson, C., Henriksen, L., Dickinson, D. C. and Levine D. (1997) 'The Early Use of Alcohol and Tobacco: Its Relation to Children's Competence and Parents' Behavior', *American Journal of Public Health,* 87(3), 359–364.

Jones, A.M. (1989) 'A Double-Hurdle Model of Cigarette Consumption', *Journal of Applied Econometrics*, vol. 4, pp. 23–39.

Kenkel, D., Lillard, D. and Mathios, A. (2006) 'The Roles of High School Completion and GED Receipt in Smoking and Obesity', *Journal of Labor Economics*, vol. 24 (3), pp. 635–660.
Kripalani, S., Paasche-Orlow, M., Parker, R., and Saha, S. (2006) 'Advancing the Field of Health Literacy', *Journal of General Internal Medicine*, 21(8), pp. 804–805.

Leitch Review of Skills (2006) *Prosperity for all in the global economy – world class skills.* London: HM Treasury.

Leigh, J. (1998): 'Parents' schooling and the correlation between education and frailty', *Economics of Education Review,* 17, pp. 349-358.

Loureiro, M.L., Sanz-de-Galdeano, A. and Vuri, D. (2006) *Smoking Habits: Like Father, Like Son, Like Mother, Like Daughter*, IZA Discussion Papers N. 2279. Bonn: Institute for the Study of Labor (IZA).

Mancuso, C. and Melina Rincon, M.A. (2006) 'Impact of Health Literacy on Longitudinal Asthma Outcomes', *Journal of General Internal Medicine*, vol. 21(8), 813–817.

Marmot, M. (2004): *Status Syndrome: How Your Social Standing Directly Affects Your Health*, p.312, Bloomsbury Publishing.

Muir, K., Santiago-Turla, C., Stinnett, S., Herndon, L., Allingham, R., Challa P. and Lee, P. (2006) 'Health Literacy and Adherence to Glaucoma Therapy', *American Journal of Ophthalmology*, 142 (2) p. 223.

Office for National Statistics (2005) *The Level of Highest Qualification Held by Young People and Adults (Labour Force Survey)*. London: DfES. www.dfes.gov.uk/rsgateway/DB/SFR/s000562/SFR06-2005v2.pdf

Osborn, C., Paasche-Orlow, M., Davis, T., and Wolf, M., (2007) 'Health Literacy: An Overlooked Factor in Understanding HIV Health Disparities', *American Journal of Preventive Medicine*, 33(5), 374–378.

Parliamentary Office of Science and Technology (2005) *Postnote 244: Binge Drinking and Public Health*. Available at: www.parliament.uk/documents/upload/postpn244.pdf

Parsons, S. and Bynner, J. (2005) *Measuring basic skills for longitudinal study: The design and development of instruments for use with cohort members in the age 34 follow-up in the 1970 British Cohort Study (BCS70)*. London: NRDC. www.nrdc.org.uk/publications_details.asp?ID=17

Sander, W. (1995) 'Schooling and quitting smoking', *Review of Economics and Statistics,* vol. 77, 191–199.

Sander, W. (1998) 'The effects of schooling and cognitive ability on smoking and marijuana use by young adults', *Economics of Education Review*, vol. 17 (3) pp. 317–324.

Sander, W. (1999) 'Cognitive ability, schooling, and the demand for alcohol by young adults', *Education Economics*, 7, pp. 53–66.

Sentell, T.L. and Halpin, H.A. (2006) 'Importance of adult literacy in understanding health disparities', *Journal of General Internal Medicine,* vol. 21, 862–866.

Shetty, N., Hartzema1, A., Bautista, R. and Glen, T. (2007) 'The association between health literacy and outcomes of care among epilepsy patients', *Research in Social and Administrative Pharmacy*, 3 (3) e7–e8.

Smith, D. and Brame, R. (2003) 'Tobit Models in Social Science Research Some Limitations and a More General Alternative', *Sociological methods and Research*, vol. 31 (3), pp. 364–388.

Vella, F. (1998) 'Estimating Models with Sample Selection Bias: A survey', *Journal of Human Resources*, vol. 33, No. 1 (Winter, 1998), pp. 127–169.

Wagstaff, A. (1986) 'The demand for health: theory and application', *Journal of Epidemiology and Community Health,* 40, pp. 1–11.

Weiss, B. D., Hart, G., McGee, D. L. and D'Estelle, S. (1992) 'Health status of illiterate adults: Relation between literacy and health status among person with low literacy skills', *Journal of the American Board of Family Practitioners*, 5, pp. 257–264.

Wolf, M. S., Gazmararian, J. A., and Baker, D. W. (2007) 'Health literacy and health risk behaviours among older adults', *American Journal of Preventative Medicine*, vol. 32 (1) 19–24.

Appendix A: Measure of alcohol consumption at age 34

Those who answered they drank ‘on most days’ were given a frequency of five days a week; those who reported drinking ‘two to three days a week’ were given a frequency of 2.5; those who reported drinking ‘once a week’, ‘two to three times a month’ or ‘less often/only on special occasions’ were given a frequency of one (their weekly alcohol consumption was likely a reflection of alcohol consumption in one sitting) and those who reported ‘never nowadays’ or ‘never drink alcohol’ were given a frequency of zero, reflecting abstinence from alcohol.

Appendix B: Descriptive statistics by quintiles

	N	Mean	Std Dev.	Min.	Max.
Daily alcohol consumption					
1 quintile	1553	1.90	0.42	0.4	2.4
2 quintile	1860	3.46	0.47	2.6	4
3 quintile	1447	5.22	0.57	4.2	6
4 quintile	1281	7.69	0.87	6.2	9.4
5 quintile	1534	14.85	5.01	9.6	32.8
Cigarettes per day					
1 quintile	903	7.97	2.58	1	10
2 quintile	102	11.90	0.30	11	12
3 quintile	499	14.89	0.44	13	15
4 quintile	662	19.84	0.65	16	20
5 quintile	200	30.66	8.21	22	80
BMI					
1 quintile	1881	20.46	1.27	9.58	21.95
2 quintile	1865	23.06	0.60	21.97	24.03
3 quintile	1879	25.08	0.61	24.04	26.11
4 quintile	1860	27.49	0.84	26.12	29.12
5 quintile	1870	33.39	4.63	29.12	89.17

Note: N = number of observations

Appendix C: Construction of literacy and numeracy measures

For the Principal Components Analysis (PCA), 47 multiple-choice individual items, coded dichotomously ('0' was incorrect; '1' was correct) were used; 27 of these items measured literacy and the remaining 20 items measured numeracy. Eigen values under 1 and factor loadings below 0.10 were dropped from the final results and varimax rotation was used to maximise the distance between the rotated factors (i.e. to make them as separate and distinct as possible).

From the unforced factor solution, it appeared there were three robust factors: two literacy factors (first two factors pulled out) and a numeracy factor. As the aim of the PCA was to pull out a 'basic skills' factor – which none of the unforced factors seemed to be – a three factor solution was forced (i.e. the assessment items were only allowed to load on to one of three factors). Again, varimax rotation was used, and eigen values below one were dropped as well as factor loadings lower than 0.10. The forced three factor solution explained 36 per cent of the variance in the data, with the first two factors explaining 26 per cent of the variance. These first two factors were used to create a measure of basic skills at age 34.

The created variable correlates in the expected direction with validity measures, such as the separate literacy and numeracy measures, education, and income. However, the most important aspect to note is that the correlations between basic skills and other variables of interest (i.e. education and income) are lower when compared with the correlation between these same variables and either literacy or numeracy.

