Disability and work: the role of health shocks and childhood circumstances

Maarten Lindeboom[a]

Ana Llena-Nozal[b]

Bas van der Klaauw[c]

October 2005

Preliminary

Abstract: In this paper we focus on the interrelation between disability and work, the role of health shocks and how this relationship varies with socio-economic background and health during (early) childhood and early adulthood. We use accidents that lead to an unscheduled hospital visit as a measure for a health shock and exploit the unanticipated nature of this health shock to assess its causal effect on disability and subsequent employment outcomes. We construct an event history model and estimate the parameters of this model on data from the British National Child Development Study (NCDS). Our empirical results show that after controlling for observed and unobserved heterogeneity employed individuals are more likely to get a health shock. Health shocks increase the rate at which individuals become non-working and disabled. Background variables like father's socio-economic status and test scores greatly influence the probability of experiencing a disability and getting out of work. The effect of a disadvantaged childhood is mainly through a strong effect on disability and employment transition rates *after* that the individuals have entered the labor market and not on the employment and disability status just after leaving full-time education.

[a] Free University Amsterdam, Tinbergen Institute, HEB, IZA, Netspar.
[b] Free University Amsterdam & Tinbergen Institute.
[c] Free University Amsterdam, Tinbergen Institute, Scholar & CEPR.

Address: Department of Economics, Free University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.
We would like to thanks Regina Riphahn, Owen O'Donell and participants of the Ecuity III workshop in Bonn, September 2005, for useful comments.
E-mail: mlindeboom@econ.vu.nl, allenanozal@econ.vu.nl, klaauw@tinbergen.nl.

1. Introduction

There exists a strong positive association between health and socioeconomic status at adulthood. Better-educated, high-income people generally have better health and lower disability rates. There are many possible mechanisms that may lead to the association in later life. Here we focus on the interaction between disability and work and the role of health shocks and socio-economic background during the (early) childhood years.

During adulthood health deteriorates with age and the rate of depreciation is influenced by decisions regarding work and life style and by shocks. Labor market choices are important because they can affect health directly and indirectly. Directly, because work or aspects of work may affect the rate at which health deteriorates. Work can have a positive effect, because of the security of a job and a stable income. It can affect the rate of depreciation negatively, if work is associated with stress and/or adverse working conditions. Employment status may also influence the likelihood of experiencing an adverse health shock. A shock or a gradual deterioration of health may lead to a long standing disability that restricts individuals in doing their daily and/or work activities and may in turn affect the individual's labor supply decisions and later work outcomes. In general it is difficult to disentangle the underlying causal mechanisms with observational data. There are unobservables that relate health and work outcomes and we need independent variation in either health or work outcomes to assess the effect of one on the other.

In this paper we use unscheduled hospital visits as a measure of a health shock and exploit the unanticipated nature of these hospital visits to identify some of the causal mechanisms. In particular, we look at the consequences of an adverse health shocks on the onset of a disability and subsequent labor market outcomes. We define a disability as a permanent chronic condition that restricts individuals in their daily activities and/or in their work. In the empirical analyses we use an event history model for the interrelation between work, disability and health shocks. To identify the causal effect of health shocks on work and disability, we require that there is unanticipated variation in the timing of health shocks. Unanticipation in this context means that the exact timing of experiencing a health shock is not known in advance to the individual. For our definition of a health shock (unscheduled hospital visits) this clearly holds. Unanticipation does not rule out that individuals may be aware that at some moments the risk that a health shock occurs is higher than in other periods or that this risk, for instance, depends on employment status. In particular, a substantial share of the adverse health shocks are related to work, which we take into account in our model framework. Also it should be stressed that we do not require health shocks to be exogenous (conditional on a set of observed characteristics). In our model we allow unobservables to affect simultaneously employment probabilities, the onset of a disability and the likelihood of experiencing a health shock. We refer to Abbring and Van den Berg (2003) for an extensive discussion on the identification of treatment effects in dynamic models.

To estimate the model, we use data from the British National Child Development Study (NCDS). The NCDS is a longitudinal study of 17,000 individuals born in Great Britain in the week of 3-9 March 1958. The data contains abundant information on the situation of the family where the individual was born in and early childhood health outcomes. The individuals are followed up to the year 2000, when they were 42 years old. At that age already about 13% of the respondents face one or more disabilities and about 17% of the disabled are out of work. These numbers show that disabilities and labor outflow are already substantial at relatively young ages.

To illustrate the mechanisms and the importance of health shocks, we perform some simulation experiments with the estimates of our model. The simulations show that health shocks are important for disability rates and that disability has a strong effect on labor outflow rates. In our simulations we also pay attention to the importance of socioeconomic background and health during early childhood. There is an extensive literature that focuses on the importance of early childhood health and economic conditions on health and socioeconomic status at adulthood (e.g. Case, Fertig and Paxson, 2005, and Currie and Hyson, 1999). We investigate whether the relation between shocks, disabilities and work at adulthood varies with socio-economic background during early childhood and conclude that it does. People from a lower socio-economic background have more often shocks, start their working career in worse health and employment outcomes and have higher probabilities of becoming disabled and out of work during their working career. The latter is due to a direct effect of early childhood circumstances and not due to an indirect effect, i.e. via the effect of work and disability status at the start of their working career[1].

The structure of the paper is as follows. Section 2 discusses the theoretical background and the empirical model. Section 3 introduces the NCDS data and reports on the variables used in the empirical part. Empirical results are discussed in Section 4. Section 5 includes some calculations and simulations. Finally, Section 6 concludes.

[1] This contrasts with what Marmot (1995) hypothesized and what was confirmed by Case et al (2005), namely that the health and labor market status just after leaving school is the most important determinant for later life health and work outcomes.

2. Theoretical background and empirical model specification

2.1 Theoretical Background

The health demand model developed by Grossman (1972) assumes that an individual inherit an initial stock of health, which depreciates with age and increases with health investments. The stock of health at a certain point in time is the accumulation of an entire history of past resources, past health behaviors and past consumption. Individuals are rational agents and according to the model they will include expectations about their health trajectories when making decisions regarding health behaviors and work. With new information, people will update their expectations and change their behavior accordingly. This underlines the difficulties in identifying the causal relations between health and socio-economic outcomes such as labor market status. If health trajectories are predictable, then individuals can anticipate to that and change their work status and other behaviors accordingly. So, in this case, an observe change in labor market status that precedes a health transition can be the results of anticipated behavior, rather then that labor market status causally affects health. Empirical analyses based on observational data are further plagued by the presence of unobserved factors that are related to both health and socio-economic status (work status).

A relatively small number of studies have used panel data and controlled for unobserved individual factors, but firm conclusions regarding the underlying causal mechanisms are not possible without reliable instruments and/or other strong assumptions regarding the interrelation between the variables of interest. A few have used natural experiments. Lindahl (2005) uses lottery prizewinners to investigate the effect of income on health. He finds a significant, but rather small effect of income on health.

The occurrence of disability can be the result of a gradual process of health deterioration, but it can also result from unforeseen health events. Smith (1998) stresses the importance of health shocks in disentangling the causal relation between health and socioeconomic status. An unforeseen shock may provide some exogenous variation in either health that is unrelated to socioeconomic status and can therefore aid in identifying the causal effect of health on socio-economic status. Smith (2003) uses the onset of chronic conditions as a measure for a health shocks and examines their effect of the probability of work, household income and wealth. He finds for a sample of individuals between 50 and 60 years old negative financial consequences of health shocks. Adams, Hurd, McFadden, Merrill and Ribeiro (2003) use hospitalization and the onset of a condition as a measure for a health shock and find some effects of health shocks on wealth. Both studies, however, do not control for unobserved factors so that the effects found in their studies may partly reflect behavioral differences that exist between those who receive and those who do not receive a health shock. Møller-Danø (2005) uses road accidents as a measure for health shocks and finds long lasting income and employment effects.

We will use unanticipated hospital visits to investigate the effect of adverse health shocks on labor market outcomes and the onset of a disability. Our data, the British NCDS data explicitly distinguishes between unanticipated events that caused hospitalization and scheduled hospitalizations. An important advantage of using this data is that in the UK health care is freely available to all individuals, which rules out selectivity in hospitalization[2]. Another important advantage is that the data follow a large cohort of individuals from birth up to age 42, which allows us to take into account much of the dynamics between shocks, the onset of a disability and work.

A dynamic model has the advantage that we can substantially relax the requirements for unanticipated hospitalizations to be valid health shocks. Smith (1998) suggests that all risk factors of experiencing the adverse health shocks should be included. Within our dynamic model we allow these health shocks to be endogenous, i.e. we allow for unobserved heterogeneity that affects both the probability of experiencing such a health shock and disability and labor market outcomes. The advantage of a dynamic model is that if health shocks are unanticipated, the effect of a health shock can be identified without exclusion restrictions or strong functional form restrictions (e.g. Abbring and Van den Berg, 2003 for an extensive discussion). We will be more specific about our dynamic model in subsection 2.2.

By now it is well documented that there is a strong association between early childhood outcomes and later life health and mortality. There are many possible explanations for the lasting influence of early childhood circumstances on health and socioeconomic outcomes during adulthood (see for an extensive summary Case, Fertig and Paxson, 2005). Poor prenatal conditions are found to be related to susceptibility to potentially life threatening diseases later in life (Barker, 1995). Poor childhood health and lower socioeconomic background may lead to worse educational outcomes and health in early adulthood, which in turn may affect later life health and socio-economic status (Marmot, 2001). Furthermore, illness at childhood may be a trigger for illnesses at adulthood. This suggests that during adulthood individuals from poor early childhood circumstances are more likely to experience adverse health shocks. Case, Fertig and Paxson, 2005, using the same data as we use in our analyses, find that childhood factors appear to operate largely through their

[2] The use of hospitalization a measure for a health shock may be problematic in the US because only a fraction (about half of the Smith's sample) is fully insured. As a consequence, the choice to go to the hospital may be related to the individual's financial situation.

effect on educational attainment and initial adult health and through a continuing direct effect of prenatal and childhood health.

In our empirical framework we allow early childhood conditions to affect health and labor market outcomes in three possible ways. First, we allow early childhood conditions to affect disability and labor market outcomes at early adulthood, which in turn may influence disability and labor market outcomes at later ages. If the hypothesis of Marmot et al (2001) is true, this would be the only relevant effect of early childhood conditions. Second, we allow for direct effects of early childhood at disability and employment status at later ages. Finally, the probability of experiencing a health shock during the course of life is allowed to depend on early childhood conditions. This implies that adverse childhood conditions may be a trigger for later health shocks, which in turn influence disability and labor market outcomes during adulthood.

2.2 Empirical specification

In this section we describe our empirical model, but before we do that we briefly sketch the structure and contents of our data. We observe individuals from birth up to the age of 42 and we are able to construct individual labor market histories since the moment of leaving full-time education. The labor market histories contain for each year whether an individual was employed or non-employed. Furthermore, for each individual we know if during the observation period the individual became disabled and if so, at which age this happened. We only focus on permanent disabilities and thus ignore short-term limitations. Finally, for each year we observe if an accident occurred to this individual. In the next section we discuss in detail the definition of our labor market states, disabilities and accidents.

The data describe the individual labor market status and health status annually. Therefore, we use a discrete-time event history model to analyze transitions between different states. The model is a semi-Markov model that contains 4 states. Let $S_l(t)$ denote the individual's labor market status at the beginning of time t, this can either be working (1) or non-working (0). In each period individuals can move between the two labor market states. Since we only follow individuals after leaving full-time education, non-working does not include full-time education. The variable $S_h(t)$ denotes the health status at the beginning of time t, which can either be disabled (1) or non-disabled (0). Because we only focus on permanent disabilities, being disabled is an absorbing state, implying that once an individual becomes disabled the individual cannot recover. The transition probabilities for moving between different states are affected by accidents that might occur to the individual. The variable $A(t)$ takes the value 1 if an accident occurred between time t and t+1 and 0 if no accident occurred in this time period. The probability of an accident is allowed to depend on the individual's current labor market status, accidents can be work related and therefore employed individuals might have higher probabilities of getting an accident. The probability that of an accident between t and t+1 equals:

$$q_t(k) = \Pr(A(t) = 1 \mid S_l(t) = k)$$

In our empirical model, we focus on the transition probabilities between the different states, which are given by

$$p_{t,(i,j),(k,m)}(a) = \Pr(S_l(t+1) = i, S_h(t+1) = j \mid S_l(t) = k, S_h(t) = m, A(t) = a)$$

Since disability is an absorbing state this transition probability equals 0 if m is disabled and j is non-disabled.

We use logit specifications to parameterize the probabilities defined above. In particular,

$$q_t(s_l(t)) = \frac{\exp(x_t\gamma + \delta s_l(t) + v_a)}{1 + \exp(x_t\gamma + \delta s_l(t) + v_a)}$$

where x_t is a vector of the individual's socioeconomic characteristics (also containing an intercept) at time t and v_a is an unobserved component that does not vary over time. The transition probabilities are specified as:

$$p_{t,(i,j),(k,m)}(a(t)) = \frac{\exp(x_t\beta_{(i,j),(k,m)} + \eta a(t) + v_{(i,j),(k,m)})}{1 + \sum_{(i',j')\neq(k,m)} \exp(x_t\beta_{(i',j'),(k,m)} + \eta a(t) + v_{(i',j'),(k,m)})}$$

if *(i,j)≠(k,m)* and

$$p_{t,(k,m),(k,m)}(a(t)) = \frac{1}{1 + \sum_{(i',j')\neq(k,m)} \exp(x_t\beta_{(i',j'),(k,m)} + \eta a(t) + v_{(i',j'),(k,m)})}$$

The transition probabilities and the probability of getting an accident are related to each other by the unobserved heterogeneity components (so v_a may be related to $v_{(i'j')(k,m)}$, $\forall$ i',j',k,m). It is well known that ignoring unobserved heterogeneity or the correlation between the different terms can cause serious biases. We use a random effects specification to model the unobserved heterogeneity, and in particular a factor-loading specification to allow for correlation between the different probabilities defined above. Define the vector w of random variables $(w_1,w_2,...,w_N)$, in which each element w_n has two discrete mass points at 0 and 1[3]. The parameter θ_k denotes the probability that the elements in w_k equals 1. The unobserved heterogeneity term follow

$$v_a = w'\alpha_a$$

and

$$v_{(i,j),(k,m)} = w'\alpha_{(i,j),(k,m)}$$

where α_a and $\alpha_{(i,j)(k,m)}$ are vectors of unknown parameters that have as many element as the vector w.

Consider an individual which we follow for T years. In this observation period the labor market states of the individual were given by $s_l(1)$, $s_l(2)$,…, $s_l(T)$ and the health states of the individual are given by $s_h(1)$, $s_h(2)$,…, $s_h(T)$ and the sequence $a(1)$,$a(2)$,…,$a(T)$ shows if an accident occurred. The likelihood contribution of this individual is given by

$$\ell = \sum_{n=1}^{N} \theta_n \left(\prod_{t=2}^{T} p_{t,(s_l(t+1),s_h(t+1)),(s_l(t),s_h(t))}(a(t)) \times q_t(s_l(t))^{a(t)} \right)$$

Note that we take the initial labor market status and health status of the individual as given. In section 4 we will estimate a multinomial logit model for these initial states, to investigate the sensitivity of the initial state to early childhood conditions.

The main parameters of interest in our model are those describing the effects of accidents on the transition probabilities. The identification of these parameters hinges on the assumption that individuals cannot anticipate the exact moment at which an accident occurs. This does not imply that an accident is exogenous or that each individual has in each time period the same probability of having an accident. The probability of having accidents can differ between individuals, based on both observed and unobserved characteristics. Furthermore, individuals

[3] The model includes an intercept and therefore the first component is normalized to 0

might know that in particular periods the probability of getting an accident is high, for example when they are employed. We only assume that in advance individuals do not know the exact timing at which an accident occurs. See Abbring and Van den Berg (2003) for an extensive discussion on identifying the effects of unanticipated interventions in dynamic models.

3. The Data

3.1 Sample

To estimate our empirical model we us the National Child Development Study (NCDS), which is a longitudinal study of about 17,000 individuals born in Great Britain in the week of 3-9 March 1958. The study started as the "Perinatal Mortality Survey" and surveyed the economic and obstetric factors associated with stillbirth and infant mortality. Since the first survey in 1958, cohort members have been traced on six other occasions to monitor their physical, educational and social circumstances. The waves were carried out in 1965 (age 7), 1969 (age 11), 1974 (age 16), 1981 (age 23), 1991 (age 33) and 1999 (age 42). In addition to the main surveys, information about the public examinations was obtained from the schools in 1978. For the birth survey, information was gathered from the mother and the medical records. For the surveys during childhood and adolescence (waves 1 to 3), interviews were carried out with parents, teachers, and the school health service; while ability tests were administered to the cohort members. The subsequent surveys included information on employment and income, health and health behavior, citizenship and values, relationships, parenting and housing, education and training of the respondents. In waves 4, 5 and 6, individuals are asked to retrospectively give information on their employment, unemployment, out-of-the-labor-force and education/training periods, recording their starting and ending dates. The NCDS is therefore highly appropriate to look at life histories and to study the impact of early life experiences on health, education and employment.

In our empirical analyses we will focus on the period in which individuals participate in the labor market. We use the waves in 1981, 1991, and 1999/2000 to construct individual labor market histories since leaving full-time education, the occurrence of accidents during adulthood and the onset of disability. To avoid the problem of left-censoring, we consider only individuals for whom we have information from the first moment of leaving full time education. Therefore, we only take into account the 12,537 individuals who participated in the 1981-survey at age 23[4]. After selecting only those with complete labor and health histories, our final sample consists of 12,448 individuals. Case, Fertig and Paxson (2005) used the same data and investigated attrition from the survey by comparing low birth weight and father's occupation across the different NCDS waves. They did not find any evidence for non-random attrition with respect to these variables. Furthermore, advisory and user support groups of the NCDS compared respondents and non-respondents in the later surveys in terms of social and economic status, education, health, housing and demography. It was found that the distribution of these variables among the sample survivors did not differ from the original sample to any great extent (NCDS User Support, 1991). In addition, the 1981 sample was compared to the UK 1981 Population Censuses in terms of the distributions of key variables such as marital status, gender, economic activity, gross weekly pay, tenure and ethnicity (Ades, 1983). The overall conclusion was that the sample appears to be representative with respect to these variables.

We performed a simple test for the presence of non-random attrition from the data by running a logit regression on participating in the 1991-wave conditional on the labor market and health status in the 1981-wave. We also included a set of individual characteristics as controls. We performed the same test for attrition from the 1999/2000-wave. The results show that attrition does depend significantly on the labor market and health status in the 1981-wave (see Tables A1 and A2). In particular, employed individuals are more likely to participate in later waves. In Subsection 4.2 we investigate the sensitivity of our parameter estimates with respect to this attrition.

The labor market status is measured each year in March. We distinguish two labor market outcomes, employed and non-employed. An individual is considered to be employed if either he has a full-time or part-time job, is self-employed or on maternity leave. Also an apprenticeship scheme which is part of a job is considered as employment. Currie and Hyson (1999), who use the same data set, show that their empirical results are not sensitive to the exact definition of employment. In Figures 1 and 2, we show for men and females at different ages the employment rate, the unemployment rate and the fraction of individuals out of the labor force and in full-time education. For men employment rates rise sharply just after the end of compulsory education at age 16. After that the fraction of employed males continues to increase until age 25, when almost everyone has left full-time education. The fraction of males out of the labor force slowly

[4] 60% of the individuals in our sample are present in wave 4 (age 23), 5 (age 33) and 6 (age 42), 28% only in wave 4 and 12% in waves 4 and 5. For these groups we also observe information on early childhood outcomes (Wave 1 and 2)

increases with age. The unemployment rate is relatively constant except for the ages 22 until 24, when there seems to be some increased unemployment. This might either be related to a business cycle effect, i.e. the recession in the late 1970s/beginning 1980s or to an age effect, i.e. youth unemployment. For the unemployment rate and the fraction of individuals in full-time education we see for females a similar pattern as for men. However, the fraction of females who is out of the labor force is much higher than for males. This fraction increases until age 28. Afterwards, when the fraction of females out of the labor force starts to decrease, employment rates increase.

In the empirical analyses we are interested in permanent disabilities or longstanding illnesses which limit an individual in his daily activities and/or work. These include, for instance, serious disability such as epilepsy, blindness, deafness, multiple sclerosis, mental retardation, a congenital condition, or a traumatic amputation or internal injury. In Appendix A we provide a list of illnesses and disorders which we consider as being permanent and limiting. This classification of disabilities coincides with the International Classification of Diseases (ICD-9) produced by the World Health Organization (1977). The ICD is extensively used in epidemiological and health management studies to classify diseases and health problems (World Health Organization, 2004). Case, Fertig and Paxson (2005), who use self-reported measures for health as outcome variable, report that these measures are very strongly correlated to chronical conditions and disabilities. Bajekal, et al (2004) show in a report commissioned by the UK Department for Work and Pensions that age-specific disability for employed workers rates do not vary much across surveys using different definitions for disability.

Figure 3 shows the fraction of individuals with a disability after age 16. Disability rates are very similar for men and women. At age 16 around 4% of the individuals in the sample has some disability. This increases up to about 13% at age 42. Some people already have long standing disabilities that started during childhood, but the majority of the disabilities started during working ages. In fact, the slope becomes steeper at older ages, which means that the hazard of onset of a disability becomes larger as people get older.

In this paper we define an accident as an unanticipated event after which an individual is admitted to hospital or attending a hospital outpatient or casualty department. We use the accidents as a measure for an unanticipated health shock. The survey has a separate question for in-patient admissions to a hospital or clinic for scheduled surgery or treatment. We observe both the date of the accident and the type of accident.[5] Men are much more likely to experience

[5] The questionnaire restricts the number of accidents that can be reported to 8 in the 1981-wave and 6 in the 1991 and 1999/2000-wave. In each wave only between 1 and 2 percent of the individuals actually reports this maximum.

accidents than women. In our sample, around 77% of the men had at least one accident during the observation period, while this was only about 42% for women. Multiple accidents for a single individual are frequently observed. Not only the incidence of accidents differs between men and women, but also the types of accidents differ. Note that a large share of the accidents takes place at work. This means that we have to take the labor market status of the individual into account when we specify our model for accidents. Table 1 lists the annual incidence rates for different types of accidents. For each type of accident men are much more likely to experience this accident than women. The most substantial difference in incidence rates occurs for work and sports-related accidents. Figure 4 shows that for both men and women the probability of having an accident is relatively high until the mid-twenties and drops substantially afterwards.

We use the annual labor market status and disability status to classify each individual in each year in one of four states: work and disabled (WD), non-work and disabled (NWD), work and non-disabled (WND) and non-work and non-disabled (NWND). In Figure 5 we show for different ages the fraction of individuals in each state. At every age most individuals are employed and non-disabled. At later ages the fraction of individuals being in non-work and non-disabled decreases while the fractions of individuals increase in both disabled states (either WD or NWD). Our empirical model is specified in terms on yearly transition probabilities between these four states. Table 2 provides for both men and women a summary of the yearly transitions. The table shows that there is a high degree of state dependence and individuals are much more likely to change labor market status than disability status.

3.2 Background variables

The NCDS has abundant information on the individual's health status and socio-economic background. For each individual we observe a range of variables that give information on an individual's health, cognitive ability and socioeconomic status during early childhood. In constructing the relevant background variables we follow the definitions used by Case, Fertig and Paxson (2005) and Currie and Hyson (1999). Table 3 provides sample means of the relevant variables. For many variables there is some item non-response and leaving these observations out of the analyses would considerable reduce our sample size. We therefore defined dummy variables indicating item non-response for some variables.

Low birth weight is a dummy variable for infants with a birth weight below 2500 grams. There is epidemiological evidence that low birth weight is strongly associated with infant and later life mortality (World Health Organization, 2004). Low weight at birth can be the result of either preterm birth (before 37 weeks of gestation) or restricted fetal growth. In the empirical analyses we do not make a distinction between these two categories. We also include height at age 23, as a (crude) measure for health. We create a dummy variable that indicates if the mother smoked after the fourth month of pregnancy. Smoking during pregnancy has been found to be related with cognitive deficiencies and other health problems in the medical and epidemiological literature (see for instance Blair et al, 1995; Conter et al., 1995; Naeye & Peters, 1984; Williams et al. 1998). Furthermore, we observe the mother's age at birth. Mother's age at the child's birth can influence the child's health through, for instance nutritional deficiencies if the mother is very young, or delivery complications if the mother is older, etc. In the empirical analyses we will include a polynomial in age.

The family's socio-economic status is derived from the father's social class at birth. The social class corresponds to a system used by the British Registrar General and consists of: professional, supervisory, skilled non-manual, skilled manual, semi-skilled non-manual, semi-skilled manual, and unskilled. We classify socioeconomic status as high if the father is in a professional, supervisory, skilled non-manual job; medium if the father is in skilled manual, semi-skilled non-manual; and low if the father is in a semi-skilled manual and unskilled job. Following Currie and Thomas (1999), we classify individuals whose father's information is missing by the mother's social class. In case the social classes of both parents are missing, we assign the individual to low socioeconomic status if the mother was single and to missing if both parents were present.

For each individual we observe test scores on math and social adjustment at age 7. Currie and Thomas (1999) show that test scores at the age of 7 have significant impacts on later education attainments and labor market outcomes. The math test is designed for the NCDS and assesses arithmetic ability. The score ranges from 0 to 10. The final test score is the Bristol Social Adjustment Guide, which is designed to assess child's behavior in school and at home, in particular the behavioral disturbances. The test is completed by the teacher who knows the child best.[6] Higher scores indicate higher maladjustment. The data also included information on the Southgate Reading Test. However, since including this test score did not improve our empirical analyses after the math score and Bristol Social Adjustment Guide were already included in the model specification. Therefore, we ignore the reading test score.

The education level is depicted by compiling an education variable with categories aggregated to national vocational qualification levels. We include the following categories: less

[6] The guide consists of a number of phrases, which describe a child's behavior, and which are grouped under a heading. Some of these headings correspond to particular sub-symptoms such as: unforthcomingness, withdrawal, depression, inconsequence, hostility, peer-maladaptiveness, etc. The teacher is asked to underline the sentences that best describe the child's behavior.

than O-levels, O-level equivalent, A-level equivalent, and degree equivalent. Finally, we will use the region at birth to control for geographical differences and/or differences in labor market conditions.

4. Empirical results

4.1 *Parameter estimates*

In this section we discuss our estimation results. We start with a model specification that does not include education as an explanatory variable. Education may take out the effects of early childhood outcomes for accidents and disabilities at later ages. We also estimated a model with education included and discuss the results in section 4.2. The joint model form accidents and transition rates includes unobservables. For the unobservables we take a factor loading specification. More specifically, we take a discrete bivariate distribution (w_1,w_2), where w_1, is associated with the probability of getting an accident and w_2 with the transition process. The random variables (w_1,w_2) are allowed to be related and each can take on two values. The parameters α (the factor loadings) are allowed to differ with each value of (w_1,w_2), for accidents and for the transition rates. We report on the parameters of the mixing distribution in the lower panel of the tables but do not discuss these any further[7].

Table 4a shows the logit specification for the probability of experiencing an accident. Being employed and being male raises the probability of an accident. This confirms what we already saw in Table 1. The accident probability is U-shaped in age; relatively high accident rates are observed for the young and the old. The table also shows that individual background, health and cognitive ability during childhood years are important. In particular, individuals whose mother smoked during pregnancy are more likely to suffer an accident and the probability of having an accident increases with the mother's age at birth. The parental socioeconomic status also has a significant effect on the accident rate. Early childhood conditions are thus important in explaining negative health shocks during adulthood. The height at age 23 is important, taller people have more accidents than small people. Individuals with a high math score at age 7 and who were less socially adjusted (high BSAG score) also have higher probabilities of getting an accident. It is difficult to connect a strong causal interpretation to these findings since, for

[7] It is difficult to interpret the findings because the parameters of Table 4 and 5 should be jointly considered. For the accidents for instance, the marginal distribution is characterized by θ_1 and the sum of mass-point 1 and mass-point 3 on the one hand and the sum of mass-point 2 and 4 on the other hand. How these accidents types relate to the different transition taypes depends on the papemeters of table 5.

example, the math score could also reflect occupational choice which is not taken into account. Finally, there is also some regional variation in the incidences of accidents.

Table 4b shows the parameter estimates of a multinomial logit model for the transition between the different labor market and disability states. These concern yearly transition probabilities and the reference group is staying in the same state. Of central importance is the effect of an accident. Accidents have a significant impact on all transitions probabilities. It is however difficult to interpret the results directly. For instance, the negative coefficient of -0.151 for the transition from WD→NWD implies that accidents reduce this transition probability relative to the recurrence probability. However, this does not imply that the transition from Work to Non-Work states is lowered when an accident occurs. For this transition one also has to consider the effect of an accident on the transitions from WND→WND, WND→WD, WND→NWND and WND→NWD. In the next subsection we will perform some calculations with the model to make the effects of accidents more insightful. For the other variables we briefly mention their effect relative to the recurrence rate.

Transition rates from work to non-work states are higher for females and the opposite holds for transition rates from non-work to work. Stated differently, females are more likely to exit work and less likely to enter into work. This is in line with results found in the labor supply and unemployment literature. Furthermore, women at work are more likely to become disabled than their male counterparts. After age 20 the probability of becoming disabled increases. There are no clear patterns in how age affects transitions between work and non-work states. It is important to note that since all individuals were born within the same week, we cannot distinguish true age effects from business cycle effects.

The variables describing the early childhood circumstances, parental socioeconomic status, mother smoking during pregnancy, mother's age at birth and the indicator for low birth weight, all have significant effects on almost all transition probabilities. In particular, more adverse early childhood conditions increase the probability of becoming disabled, the incidence of entering non-employment and the length of non-employment spells. Early childhood conditions thus have a significant direct effect on the rate of health depreciation and changes in employment rates over the life cycle.

Individuals with a high math score at age 7 and who were more socially adjusted are significantly less likely to become disabled and non-employed (high BSAG scores are associated with lower social adjustment). When non-employed, these individuals have higher transition rates and hence on average short non-employment spells. Relatively tall people at age 23 are more likely to become disabled than shorter people. When we condition on disability status, we see that when tall people are non disabled that they are much more likely to be employed, i.e. they have a significant lower transition probability from employment to non-employment and a significant higher transition probability from non-employment to employment. Furthermore, there is some significant regional variation in transition probabilities.

4.2 *Sensitivity analyses*

Currie and Hyson (1999) investigate the effects of early childhood conditions on employment, health and wages. Their empirical results indicate that these effects actually differ between men and women. In particular, the effects of early childhood conditions are for women pronounced at younger ages than for men. Therefore, we also estimated our model separately for males and females. The results are reported in Tables 5a and 5b. For accidents most of the effects remain qualitatively the same. For the transition rates we observe some qualitative differences. The most important differences concern the transitions between work states for disabled workers. From the table it is difficult to judge whether these differences are quantitatively important. Simulations with the different models may shed some more light on this. We return to this in the next section.

In the previous subsection we did not include the individual's education level as regressor. In Subsection 2.1 we argued that education can be an intermediate variable for early childhood conditions. For example, Currie and Hyson (1999) show that the effects of early childhood conditions are largest on educational attainments. However, education is also a proxy variable for occupation and human capital. Therefore, it is likely to have a substantial effect on labor market and health outcomes (Fuchs, 2004). If we want to test if early childhood circumstances have an effect on the rate of health depreciation, we should include the level of education in the model. Furthermore, estimating the model with the level of education as regressor provides an indication on the robustness of the effects of accidents.

Table 6 shows the estimation results for a model specification with the level of education included. The education coefficients are significant, but there are no important changes in the magnitude and significance of the other variables once education is included. The reference group for education is those with an education below O-levels. Having higher education appears to decrease the probabilities of becoming disabled and non-employed. The likelihood of an accident is only higher for those with A-levels.

In Section 3 we showed that attrition from the panel is non-random with respect to the labor market and disability status at age 23. To check if the attrition has an effect on the main conclusions from the model, we estimate the model again but with a dummy variable indicating if the individual drops out of the panel before the final wave. In Table 7 we show the estimation results including this additional dummy variable

(Table 7: TO BE ADDED).

5. Simulations with the model

5.1 *Initial state*

In this section we perform some simulation experiments to investigate the importance of health shocks on disability and labor market transitions and to get some insight on how important early childhood conditions are on outcomes during adulthood. In particular, we want to get some insight into the importance of the different mechanisms through which early childhood conditions work. However, in our model we took the initial state of each individual as given. Therefore, before presenting the results from the simulation experiments, we first estimate a model for the initial state.

Table 8 shows the estimation results for a multinomial logit model for the initial state, which is the first state after leaving full-time education. Compared to the earlier estimations, we did not include age a regressor as there is only little variation in the age at which individuals leave full-time education and it appeared to have little explanatory power. Women are less likely to be disabled than men. Taller people and individuals with a high math score at age 7 and those who are more socially adjusted have significantly higher probabilities of being employed and non-disabled after leaving school. The variables describing early childhood conditions most often do not have a significant impact; only individuals whose mother smoked during pregnancy and had parents from a low socioeconomic background are significantly more likely to be non-employed and non-disabled. This hints that Marmot's the pathways' hypothesis (see Marmot et al, 1999), i.e. that the effects of early childhood conditions on later age health and socio-economic status mainly works via its effect on health and socioeconomic status at early adulthood.

5.2 Simulations

The model estimates can be used to perform some simulations that give us more insight into the importance of health shocks as measured by our accidents variable and background variables on transition rates between disability and employment states. We use Tables 4aa, 4b and 8 for the simulations.

Figure 6 depicts work and disability probabilities for each individual in the sample, starting at age 16. Hence, this is an informal check on the fit of the model. Figure 6 compares rather well with the observed probabilities depicted in Figure 5.

We first investigate the effect of accidents on the probability of getting a disability. Next, we look at the impact of a disability on employment rates. Finally, we look at the role of childhood conditions on both disability and employment. The results are shown in Tables 9, 10 and 11 respectively.

We look at the effect of accidents through different scenarios: (a) getting an accident every year, (b) no accident and (c) one accident at age 25. In Figure 7 we compare the predicted disability rates of the model with the simulated disability rates in the different scenarios. If everyone would have an accident every year, disability rates at age 40 would be twice as high as in case no-one would ever have an accident. If no individual would get an accident disability rates at age 24 are 4.8%, this is only slightly lower than the average in the sample (4.9%). An accident at age 25 increases the disability rate in the next year with 8% (from 4.9% to 5.3%). Disability is an absorbing state and hence disability rates remain high after the single health shock (accident).

We found that the incidence of an accident is higher for those at work we and therefore considered the following little experiment. We first consider the case where the incidence of adverse health shocks is reduced with 25% and next we consider the case where work does not affect the incidence of health shocks. The latter experiment could for instance mimic the effect of a workplace safety policy or a policy aimed at reducing work stress. The (indirect) effects of these experiments on disability and employment rates are very small. For instance, the disability rate at age 40 is reduced from 11.8% to 11.6%, while employment rates in both experiments go up from 88.1% to 88.2%. We can see that reducing total accidents by 25% is almost equivalent to removing the effect of work on accidents.

For the effect of disability on employment rates we compare the following cases: what are the employment rates given (a) the disability rates predicted by the model (b) when nobody is ever disabled, and (c) when everybody becomes disabled at age 25. From figure 8 it can be seen that at age 25, employment rates drop drastically (by 2.5%) if individuals get a disability and this decline in employment continues over time. The gap between employment rates widens for the different disability scenarios. By age 40 the difference in employment is of 19.8% between the non-disabled and the disabled case (see Table 10).

Finally, we explore the role of childhood conditions and assume that everybody comes from a high socio-economic background (labeled as high SES). We in addition assume that other background variables are favorable (i.e. no low birth-weight and no maternal smoking). We simulate the effect of SES on (a) initial state, (b) direct transition probabilities, and (c) accident probabilities. We do this to see where early childhood conditions are most important. Tables 11 and 12, present the effects of childhood characteristics on disability and employment (respectively) at different stages. The first observation is that childhood characteristics appear to matter more in the transition probabilities and have the less impact on the probability of experiencing accidents (see Figure 9). Indeed, it seems that coming from a high SES does not greatly affect the rate of health shocks in order to substantially reduce disability. Likewise, high SES has only a limited impact on the probability of being disabled after completing full-time education, the initial state (11.7% versus 11.8%). On the other hand, disability rates in adulthood are greatly reduced (10%). This is because the socio-economic background continues to influence the transition probabilities. The effects on employment rates are very similar. High SES increases employment only through its effects on the transition probabilities (91.4% versus 88.1%). This can be shown in Figure 10, where only the line for the high SES during the transitions differs from the predicted by the model.

For comparison purposes, we perform the same simulations with low SES, that is, we look at how employment and disability rates vary when everybody comes from low SES (and all mothers smoked during pregnancy and all had low birth weight) during the three different pathways: (a) initial state, (b) direct transition probabilities, and (c) accident probabilities. It can be seen from figure 11 that employment rates at age 40 in cases (a) and (c) are quite similar to the rates predicted by the model (87.8%, 88.1% and 88.1% respectively). This again demonstrates the limited impact of SES on initial employment conditions and the rate of health shocks. SES is, on the other hand, important for its employment effects during the transitions in adulthood. Indeed, if all individuals were from a low SES employment would be reduced to 81.3%. Figure 12 shows the gap in the evolution of employment for those of low and high SES during the transitions. By age 40, this gap is of 11.1%.

The effect of SES on disability through the different paths is somehow different. This is related to the fact that SES appears to have a somehow more pronounced effect on the initial disability rates after finishing full-time education. In figure 13 we can see that if people have a low SES during the initial status, disability rates at age 16 would be much higher than in the other scenarios (6.7% versus 4.5%). Nevertheless, when individuals are from a low SES during the transitions, the disability rate increases much faster and, by age 40, it is higher than in any other scenario (15.8% compared with 13.6% for the initial state scenario and 11.8% for both the accident scenario and the predicted probability). The difference in disability rates for the effects of SES on transitions is of almost 6 percentage points at age 40 between those from high and those from low SES, as can be seen in figure 14.

6. Conclusion

This paper explores the relationship of disability and work over the life cycle. We are particularly interested in the effect of a health shock on later employment and disability outcomes and we want to examine whether this relationship differs with socio-economic conditions during early childhood. We use unanticipated hospital admissions (labeled as accidents) as a measure of a health shock. In a dynamic model, we exploit the unanticipated nature of our health variable to assess its causal effect on disability and labor market status. We estimate our model on data from the British National Child Development Study (NCDS).The results indicate that indeed current labor market status greatly increases the probability of experiencing health shocks (by 40%). Accidents have a strong impact on the individual's outcomes; in particular, the occurrence of a disability is more than twice as likely to happen after experiencing an accident. Furthermore, individuals with disabilities have a much higher probability of entering unemployment. Finally, early childhood circumstances have a direct effect on becoming disabled and non-employed during the course of life. Indeed, individuals whose parents where from low SES have 40% more chances of becoming disabled and 22% more chances to loose their job due to disability.

Our results are particularly relevant for policy matters as we postulated in the introduction. They are partly in line with previous literature findings (Case, Lubotski & Paxson, 2002; Case, Fertig & Paxson, 2005; Currie & Hyson, 1999) where it has been found that lower income children are at higher risk of worse health, and that the effects are long lasting. According to the pathways models, childhood circumstances do not affect adult health risk directly but indirectly through its effects on adult social circumstances. Our findings suggest, that parental low SES has a limited effect on health outcomes at early adulthood and a much stronger effect on the likelihood of being disabled later in adulthood. This is the case because the parental socio-economic background appears to have a strong effect on disability and employment transition rates *after* that the individuals have entered the labor market. We find that socio-economic background has a significant, but quantitatively small effect on the accident rates. Therefore the indirect effect of low socio-economic status, via the occurrence of a health shock, is very small. From this one can conclude that the larger part of the effect of low socio-economic status during early life comes from an accumulation of higher transition rates to disability and non-employment.

This conclusion is important for public policy since it implies that a policy that improves early childhood outcomes for the economically disadvantaged will reduce the odds of experiencing a permanent disability later in life. This in turn will positively affect the work patterns of workers later in life. Policies aimed at the young can thus positively influence health and work outcomes at advanced ages.

References

Abbring JH, van den Berg GJ. The nonparametric identification of treatment effects in duration models. *Econometrica 2003;* **71**: 1491-1517.

Adams, P, Hurd MD, McFadden D,Merrill A, Ribeiro T. Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status, *Journal of Econometrics* 2003; **112**: 3-56.

Ades, T. Comparing NCDS4 to the 1981 UK Census. *National Child Development Studies Working Paper Series,* Fourth Follow-up Working Paper 11, 1983.

Anderson KH, Burkhauser R. The Retirement-Health Nexus: A New Measure of an Old Puzzle. *Journal of Human Resources* 1985; 2**0(3)**: 315-330.

Bajekal M, Harries T, Breman R, Woodfield K. Review of disability estimates and definitions. *IAD Social Research Division In-house Reports*, In-House Report No. 128, 2004.

Barker DJP. Fetal origins of coronary heart disease. *British Medical Journal* 1995; **311**: 171–174.

Blair, PS, Fleming PJ, Bensley D, Smith I, Bacon C, Taylor E, gloding J, Tripp J. Smoking and the sudden infant death syndrome: results from 1993-5 case-control study for confidential inquiry into stillbirths and deaths in infancy. *British Medical Journal* 1996; **313**: 195 - 198.

Bound J, Burkhauser RV. Economic Analysis of Transfer Programs Targeted on People with Disabilities, In *Handbook of Labor Economics*, Ashenfelter OC, Card D (eds.). Amsterdam: Elsevier Science, 1999; Vol. 3C, 3417-3528.

Bound J, Schoenbaum M, Stinebrickner TR, Waidman W. The dynamic effects of health on the labor force transitions of older workers. *NBER Working Papers Series*, Working Paper 6777, 1998.

Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. *Journal of Health Economics* 24, 365-389, 2005.

Case A, Lubotsky M, Paxson C. Economic Status and Health in Childhood: The Origins of the Gradient, *American Economic Review*, 2002; **92(5):** 1308-1334.

Conter V, Cortinovis I, rogari P, Riva L. Weight growth in infants born to mothers who smoked during pregnancy. *British Medical Journal*, Mar 1995; **310**: 768 - 771.

Currie J, Hyson R.Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight. *American Economic Review Papers and Proceedings* 1999; **89(2)**: 245-250.

Currie, J., and B.C. Madrian, 1999. Health insurance and the labour market, in *Handbook of Labor Economics*. Ashenfelter, O., Card, D., eds., Volume 3C, ch.50.

Currie J, Stabile M. Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children? *American Economic Review* 2003; **93(5):** 1813-1823.

Currie J, Thomas D. Early Test Scores, Socioeconomic Status and Future Outcomes. *NBER Working Papers Series*, Working Paper 6943, 1999.

Doblhammer G. The late life legacy of very early life. *MPIDR*, Working Paper 2003-03, 2003.

Fuchs VR. Reflections on the socio-economic correlates of health. *Journal of Health Economics* 2004; **23(4):** 653-661.

Grossman M. On the concept of health capital and the demand for health. *Journal of Political Economy* 1972; **80:** 223-255.

Haveman, R. Wolfe, B., Kreider B., and M. Stone, 1994. Market work, wages, and men's health. *Journal of Health Economics* 13, 163-182.

Kerkhofs M, Lindeboom M. Age related health dynamics and changes in labour market status. *Health Economics* 1997; **6**: 407-423.

Lindahl, M. Estimating the effect of income on health and mortality using lottery prizes as exogenous source of variation, *Journal of Human Resources* 2005; **40**: 144-168.

Lumsdaine R, Mitchell O. New Developments in the Economic Analysis of Retirement, In *Handbook of Labor Economics*, Ashenfelter OC, Card D (eds.). Amsterdam: Elsevier Science, 1999; Vol. 3.

Marmot M, Shipley M, Brunner E, Hemingway H. Relative contribution of early life and adult socioeconomic factors to adult morbidity in the Whitehall II Study. *Journal of Epidemiology and Community Health* 2001; **55**: 301-307.

Naeye RL , Peters EC. Mental development of children whose mothers smoked during pregnancy. *Obstetric Gynecoogy* Nov 1984; **64**: 601 - 607.

NCDS User Support Group. *NCDS5 Report*. City University: London, 1991.

Sickles RC, Taubman P. An analysis of the health and retirement status of the elderly. *National Bureau Economic Research Working Paper Series*, Working Paper 1459, September 1984.

Sickles R, Yazbeck AS. On the Dynamics of Demand for Leisure and Production of Health: Evidence from the Retirement History Survey. *Journal of Business and Economic Statistics* 1998: 187-197.

Smith, JP. Socioeconomic status and health. *American Economic Review Papers and Proceedings* 1998; **88:** 192-196.

Smith, JP. Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives* 1999; **13(2):** 145-166.

Smith, JP. Consequences and Predictors of New Health Events. *National Bureau Economic Research Working Paper Series*, Working Paper 10063, November 2003.

Smith A, Twomey B. Labour market experiences of people with disabilities. *Labour Market Trends*, August 2002.

UNICEF & WHO. *Low birthweight*. Geneva: WHO Publications, 2004.

Williams, GM, O'Callaghan M, Najman JM, Bor W, Andersen MJ, Richards D, Chinlyn U. Maternal Cigarette Smoking and Child Psychiatric Morbidity: A Longitudinal Study. *Pediatrics*, Jul 1998; **102**: 11.

Table 1: Yearly incidences of different types of accidents

	Male	Female
Road (pedestrian)	0.0018	0.0013
Road (driver)	0.0179	0.0080
Workplace	0.0398	0.0072
Home	0.0127	0.0107
Sports	0.0338	0.0047
Other	0.0139	0.0072

Table 2: Transition matrices for work and disability states by gender

Male					Female				
	WD(t)	NWD(t)	WND(t)	NWND(t)		WD(t)	NWD(t)	WND(t)	NWND(t)
WD(t-1)	95.3%	4.7%			WD (t-1)	90.3%	9.7%		
NWD(t-1)	16.8%	83.2%			NWD (t-1)	12.8%	87.2%		
WND(t-1)	0.3%	0.1%	96.8%	2.8%	WND(t-1)	0.3%	0.0%	91.7%	7.9%
NWND(t-1)	0.3%	0.7%	41.9%	57.2%	NWND(t-1)	0.1%	0.4%	19.3%	80.2%

Table 3: Sample mean of the individual characteristics			
	Total	Male	Female
Female	50.1%		
Parental socioeconomic status at birth			
Missing	6.3%	6.6%	6.0%
High	25.6%	25.9%	25.3%
Medium	47.1%	46.5%	47.7%
Low	21.0%	21.0%	21.0%
Mother smoked after the fourth month of pregnancy			
Missing	6.3%	6.5%	6.1%
Yes	30.8%	30.3%	31.3%
No	62.9%	63.1%	62.6%
Mother's age at birth (in years)	27.6	27.6	27.6
Missing	5.2%	5.4%	4.9%
Height at age 23 (in meters)	1.70	1.77	1.62
Missing	0.7%	0.7%	0.6%
Birth weight			
Missing	5.5%	5.8%	5.2%
Low (less than 2500 grams)	4.8%	4.1%	5.4%
Normal (more than 2500 grams)	89.7%	90.1%	89.3%
Math test score at age 7 (scale 0-10)	5.1	5.1	5.0
Missing	11.3%	11.9%	10.8%
Bristol Social Adjustment Guide at age 7	8.3	9.7	6.9
Missing	11.2%	11.8%	10.7%
Region of residence at birth			
Missing	5.1%	5.4%	4.9%
North	27.2%	26.6%	27.8%
Midlands	23.5%	24.3%	22.7%
South & Wales	16.4%	16.2%	16.5%
Scotland	10.5%	10.2%	10.8%
London & South-East	17.4%	17.4%	17.3%
Education (National Vocational Qualification level)			
Below O-levels equivalent	26.1%	24.5%	27.7%
O-level equivalent	31.4%	27.7%	35.0%
A-level equivalent	17.0%	20.8%	13.3%
Degree equivalent	25.6%	27.1%	24.1%

Table 4a: Logit for probability of experiencing an accident

	Parameter estimate	Standard error
Intercept	0.120	0.007
Being employed	0.371	0.009
Female	-1.036	0.009
Age (divided by 10)	-1.686	0.003
Age squared (divided by 100)	0.221	0.002
Parental socioeconomic status at birth		
missing	0.041	0.004
high	-0.063	0.007
low	-0.047	0.006
Mother smoked at pregnancy	0.089	0.007
missing	0.224	0.004
Mother’s age at birth		
age (divided by 10)	-0.468	0.004
age squared (divided by 100)	0.720	0.004
missing	-1.074	0.003
Height at age 23	1.220	0.007
missing	1.900	0.004
Low birth weight	0.005	0.004
missing	-0.218	0.003
Math score at age 7	0.112	0.009
missing	0.046	0.006
Bristol Social Adjustment Guide at age 7	0.764	0.012
missing	-0.082	0.008
Region of residence at birth		
missing	0.217	0.005
North	0.047	0.008
Midlands	0	
South & Wales	0.024	0.004
Scotland	-0.105	0.005
London & South-East	0.035	0.004
Unobserved heterogeneity (factor loading)		
Probability 1: $\theta_1\ \theta_2$	0.162	0.0004
Probability 2: $(1-\theta_1)\theta_2$	0.104	0.0003
Probability 3: $\theta_1\ (1-\theta_2)$	0.447	0.0012
Probability 4: $(1-\theta_1)(1-\theta_2)$	0.287	0.0008
Location mass point 1	0	
Location mass point 2	1.190	0.005
Location mass point 3	-0.984	0.007
Location mass point 4	0.206	0.004

Table 4b: Multinomial logit with unobserved heterogeneity on transitions between work and disability states

	WD to NWD	NWD to WD	WND to WD	WND to NWD	WND to NWND	NWND to WD	NWND to NWD	NWND to WND
Intercept	-2.321	-2.111	-6.797	-7.242	-3.402	-3.159	-3.768	3.082
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)	(0.004)	(0.005)	(0.010)
Accidents	-0.151	0.154	0.816	1.444	0.064	0.739	0.864	0.190
	(0.003)	(0.003)	(0.003)	(0.003)	(0.010)	(0.011)	(0.003)	(0.005)
Female	0.794	-0.447	0.294	0.864	0.961	-1.184	-0.689	-0.867
	(0.008)	(0.034)	(0.005)	(0.003)	(0.008)	(0.004)	(0.004)	(0.005)
Age (divided by 10)	-0.180	0.645	-0.111	-1.429	2.610	-0.764	-1.035	-2.244
	(0.004)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.003)	(0.006)
Age squared (divided by 100)	-0.039	-0.161	0.106	0.314	-0.598	0.105	0.246	0.340
	(0.004)	(0.004)	(0.004)	(0.005)	(0.003)	(0.012)	(0.005)	(0.004)
Parental socioeconomic status at birth								
missing	0.199	-0.066	0.130	0.172	0.282	0.135	-0.408	-0.153
	(0.004)	(0.004)	(0.006)	(0.007)	(0.003)	(0.003)	(0.011)	(0.003)
high	-0.198	-0.102	-0.184	-0.511	-0.157	0.375	-0.165	0.210
	(0.003)	(0.011)	(0.007)	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)
low	0.223	-0.126	0.187	0.215	0.246	0.456	-0.174	-0.146
	(0.003)	(0.017)	(0.007)	(0.003)	(0.006)	(0.007)	(0.003)	(0.004)
Mother smoking at pregnancy	0.158	-0.017	0.191	0.395	0.179	-0.167	0.172	-0.040
	(0.004)	(0.009)	(0.006)	(0.003)	(0.007)	(0.004)	(0.003)	(0.006)
missing	0.703	-0.501	0.073	0.540	0.221	-0.284	-0.363	-0.204
	(0.003)	(0.004)	(0.006)	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)
Mother's age at birth								
age (divided by 10)	0.081	-0.355	-0.469	0.192	-0.282	0.094	-0.114	-0.107
	(0.004)	(0.012)	(0.005)	(0.003)	(0.006)	(0.005)	(0.003)	(0.007)
age squared (divided by 100)	0.133	0.303	0.739	-0.511	0.444	0.132	0.282	0.281
	(0.003)	(0.013)	(0.004)	(0.005)	(0.008)	(0.004)	(0.004)	(0.005)
missing	-0.150	0.025	-0.331	-0.553	-0.378	0.013	0.140	-0.082
	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)	(0.003)	(0.003)	(0.006)
Height at 23	0.013	0.600	0.519	0.419	-1.339	0.007	0.518	0.873
	(0.004)	(0.013)	(0.008)	(0.008)	(0.004)	(0.008)	(0.006)	(0.009)
missing	-0.385	0.249	-0.475	0.272	-1.968	0.134	-0.401	1.207
	(0.003)	(0.005)	(0.005)	(0.005)	(0.019)	(0.004)	(0.003)	(0.010)
LBW	0.048	-0.456	0.206	-0.261	-0.068	0.177	-0.052	-0.099
	(0.003)	(0.008)	(0.003)	(0.007)	(0.019)	(0.003)	(0.015)	(0.013)
missing	0.062	-0.267	-0.309	-0.344	-0.196	0.590	0.069	-0.043
	(0.008)	(0.006)	(0.012)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
Math score at age 7	-0.103	0.056	-0.065	-0.031	-0.527	0.001	-0.032	0.363
	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)	(0.003)	(0.003)	(0.007)
missing	0.101	-0.955	0.168	0.038	0.021	-0.456	0.376	0.076
	(0.004)	(0.033)	(0.027)	(0.007)	(0.008)	(0.00.)	(0.003)	(0.010)
BSAG at age 7	0.472	-0.410	0.144	0.061	3.269	-0.048	0.104	-2.046
	(0.005)	(0.005)	(0.004)	(0.003)	(0.038)	(0.003)	(0.003)	(0.027)

missing	-0.176	0.651	-0.085	-0.336	0.071	-0.100	-0.319	-0.046
	(0.011)	(0.022)	(0.027)	(0.005)	(0.009)	(0.005)	(0.004)	(0.011)
Region of residence at birth								
Missing	-0.059	-0.031	-0.181	-0.331	0.163	0.041	0.109	0.495
	(0.004)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.006)
North	0.395	-0.119	-0.077	0.192	0.163	0.116	0.415	-0.015
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.006)
South/Wales	0.170	-0.140	0.180	0.122	0.037	0.078	0.261	-0.020
	(0.004)	(0.006)	(0.003)	(0.004)	(0.005)	(0.011)	(0.005)	(0.006)
Scotland	0.199	0.013	-0.014	-0.293	0.120	0.357	0.394	-0.025
	(0.003)	(0.006)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.005)
London	0.012	-0.127	-0.130	-0.284	0.012	-0.347	0.359	-0.004
	(0.003)	(0.004)	(0.003)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Location mass point 1	0	0	0	0	0	0	0	0
Location mass point 2	-0.628	1.125	0.262	-0.884	-0.209	-0.814	-0.100	-0.529
	(0.004)	(0.004)	(0.003)	(0.009)	(0.004)	(0.014)	(0.005)	(0.006)
Location mass point 3	-1.429	0.361	-0.546	-1.694	-1.102	-0.386	-0.634	-0.658
	(0.008)	(0.007)	(0.004)	(0.005)	(0.006)	(0.004)	(0.003)	(0.014)
Location mass point 4	-2.057	1.486	-0.284	-2.578	-1.932	-1.200	-0.734	-1.187
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Value of the log-likelihood	-105,348.420							

Table 5a: Logit for probability of experiencing an accident - Males

	Parameter estimate	Standard error
Intercept	-0.420	0.000
Being employed	0.387	0.000
Age (divided by 10)	-1.534	0.000
Age squared (divided by 100)	0.182	0.000
Parental socioeconomic status at birth		
missing	0.080	0.000
high	-0.101	0.000
low	-0.053	0.000
Mother smoked at pregnancy	0.106	0.000
missing	0.274	0.000
Mother’s age at birth		
age (divided by 10)	-0.727	0.000
age squared (divided by 100)	1.068	0.000
missing	-1.468	0.000
Height at age 23	1.059	0.000
missing	1.661	0.000

Low birth weight	-0.064	0.000
missing	-0.073	0.000
Math score at age 7	0.897	0.000
missing	0.054	0.000
Bristol Social Adjustment Guide at age 7	1.645	0.000
missing	-0.150	0.000
Region of residence at birth		
missing	-0.011	0.000
North	0.125	0.000
Midlands	0	
South & Wales	0.075	0.000
Scotland	-0.045	0.000
London & South-East	0.036	0.000
Unobserved heterogeneity (factor loading)		
Probability 1: $\theta_1\ \theta_2$	0.157	0.000
Probability 2: $(1-\theta_1)\theta_2$	0.114	0.000
Probability 3: $\theta_1\ (1-\theta_2)$	0.421	0.000
Probability 4: $(1-\theta_1)(1-\theta_2)$	0.307	0.000
Location mass point 1	0	
Location mass point 2	1.156	0.000
Location mass point 3	-0.907	0.000
Location mass point 4	0.249	0.000

Table 5a: Multinomial logit with unobserved heterogeneity on transitions between work and disability states –Males

	WD to NWD	NWD to WD	WND to WD	WND to NWD	WND to NWND	NWND to WD	NWND to NWD	NWND to WND
Intercept	0.499	-2.827	-6.057	-7.319	2.917	-3.371	-4.526	-1.675
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Accidents	-0.309	-0.133	0.834	1.930	0.070	1.987	1.062	-0.015
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age (divided by 10)	-3.460	0.959	-0.049	-4.837	1.258	-2.888	-2.117	-1.896
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age squared (divided by 100)	0.541	-0.238	0.092	0.909	-0.369	0.482	0.451	0.243
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Parental socioeconomic status at birth								
Missing	0.008	-0.714	0.196	1.129	0.615	2.417	-2.283	-0.518
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
High	-0.454	0.190	-0.223	-0.299	-0.047	0.789	0.886	0.228
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Low	0.422	-0.322	0.180	0.007	0.510	0.640	-0.244	-0.267
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother smoking	0.236	0.017	0.230	0.849	0.338	-0.528	0.209	-0.082

at pregnancy								
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.916	-1.367	0.574	-1.867	0.491	-1.320	-1.903	0.136
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's age at birth								
age (divided by 10)	1.859	-0.630	-1.053	0.790	-0.842	-0.793	-0.235	0.246
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
age squared (divided by 100)	-3.085	0.101	1.592	-1.307	1.472	1.245	-0.115	-0.556
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
missing	-0.062	0.627	-0.463	-0.922	-1.374	-0.474	0.780	-1.407
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Height at 23	0.165	0.900	0.822	3.003	-2.737	0.432	1.196	2.299
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	-0.307	1.400	-1.559	-0.206	-4.697	-0.189	-2.104	3.837
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LBW	-0.079	-0.925	0.205	0.556	0.118	-2.326	3.132	-0.307
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	2.338	-0.449	-2.093	-1.370	-0.663	0.212	-0.135	-0.336
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Math score at age 7	-0.492	0.289	-0.438	-0.398	-1.187	-0.040	-0.509	0.778
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	-1.715	-0.015	0.292	-0.464	0.051	-0.715	-0.912	-0.430
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BSAG at age 7	1.784	-1.004	1.575	-0.823	14.259	-0.358	0.440	-8.257
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	1.820	-1.393	-0.256	-0.702	0.170	-0.247	0.535	0.415
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Region of residence at birth								
Missing	-0.040	0.784	0.399	-1.272	0.069	0.194	0.479	2.459
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
North	0.688	0.060	-0.201	0.097	0.319	1.235	0.623	-0.150
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
South/Wales	0.248	-0.302	0.088	-0.056	-0.046	1.777	-0.268	0.040
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Scotland	0.383	-0.301	-0.219	-0.974	0.263	1.976	0.006	-0.184
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
London	-0.146	-0.293	-0.112	-1.097	-0.027	1.754	1.198	0.145
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Location mass point 1	0	0	0	0	0	0	0	0
Location mass point 2	-0.761	1.872	0.085	-2.509	-1.345	-0.305	0.018	1.103
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Location mass point 3	-2.014	1.565	-0.500	-1.336	-1.885	1.504	0.232	1.037
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Location mass	-2.526	3.437	-0.415	-3.845	-3.230	1.199	0.250	2.140

point 4								
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Value of the log-likelihood	-53,566.645							

Table 5b: Logit for probability of experiencing an accident - Females

	Parameter estimate	Standard error
Intercept	-1.047	0.000
Being employed	0.267	0.000
Age (divided by 10)	-2.302	0.000
Age squared (divided by 100)	0.358	0.000
Parental socioeconomic status at birth		
missing	-0.064	0.000
high	0.049	0.000
low	-0.046	0.000
Mother smoked at pregnancy	0.026	0.000
missing	0.077	0.000
Mother's age at birth		
age (divided by 10)	0.083	0.000
age squared (divided by 100)	-0.057	0.000
missing	0.595	0.000
Height at age 23	1.276	0.000
missing	1.890	0.000
Low birth weight	0.059	0.000
missing	-0.591	0.000
Math score at age 7	0.528	0.000
missing	0.017	0.000
Bristol Social Adjustment Guide at age 7	0.144	0.000
missing	0.033	0.000
Region of residence at birth		
missing	-0.005	0.000
North	-0.108	0.000
Midlands	0	
South & Wales	-0.132	0.000
Scotland	-0.224	0.000
London & South-East	0.012	0.000
Unobserved heterogeneity (factor loading)		
Probability 1: $\theta_1\ \theta_2$	0.128	0.000
Probability 2: $(1-\theta_1)\theta_2$	0.308	0.000
Probability 3: $\theta_1\ (1-\theta_2)$	0.166	0.000
Probability 4: $(1-\theta_1)(1-\theta_2)$	0.398	0.000
Location mass point 1	0	
Location mass point 2	-1.466	0.000
Location mass point 3	-0.700	0.000
Location mass point 4	-2.165	0.000

Table 5b: Multinomial logit with unobserved heterogeneity on transitions between work and disability states –Females

	WD to NWD	NWD to WD	WND to WD	WND to NWD	WND to NWND	NWND to WD	NWND to NWD	NWND to WND
Intercept	-0.585	-2.034	-4.526	-1.680	-3.512	-1.784	-1.106	0.743
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Accidents	0.330	0.431	0.799	0.992	-0.050	-1.223	0.624	0.250
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age (divided by 10)	0.976	1.096	0.316	-3.545	3.496	-3.267	-2.552	-1.964
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age squared (divided by 100)	-0.240	-0.225	0.033	0.664	-0.744	0.510	0.502	0.317
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Parental socioeconomic status at birth								
Missing	0.345	-0.693	0.248	0.861	0.254	-1.102	-1.222	0.096
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
High	-0.040	-0.089	-0.137	-0.520	-0.177	0.450	-0.738	0.185
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Low	-0.049	-0.160	0.189	0.404	0.102	0.595	-0.238	-0.120
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother smoking at pregnancy	0.119	-0.098	0.154	-0.039	0.105	-0.059	0.202	-0.026
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	-0.039	-1.325	-1.157	1.271	0.184	-1.489	-1.024	-0.391
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's age at birth								
age (divided by 10)	-1.238	-1.515	-1.716	-0.110	-0.429	-0.244	-0.743	-0.345
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
age squared (divided by 100)	2.488	2.548	-3.006	-0.350	0.676	0.502	1.408	0.744
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
missing	-0.086	0.110	-1.142	-2.284	-0.578	-0.766	-0.383	-0.623
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Height at 23	-0.773	0.735	0.203	-0.123	-1.506	-0.815	0.198	0.642
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	-1.545	0.287	-1.064	-1.242	-2.092	-0.631	-2.349	0.885
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LBW	-0.188	-0.342	0.265	-0.229	-0.101	0.500	-0.494	-0.049
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	-0.166	-0.383	-0.253	-0.554	0.022	1.164	-0.172	0.024
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Math score at age 7	-0.181	-0.007	-0.144	-0.032	-0.588	-0.033	-0.085	0.270

	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	-0.112	-0.066	0.164	-0.183	0.097	-0.466	0.396	0.223
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BSAG at age 7	0.607	-0.447	-0.130	0.055	4.065	-0.085	0.282	-2.712
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	-0.046	0.074	0.130	0.079	-0.003	0.133	-0.049	-0.219
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Region of residence at birth								
Missing	-0.692	0.061	-0.114	-0.286	-0.456	0.091	0.142	0.620
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
North	0.478	-0.239	0.021	0.033	0.068	0.074	0.216	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
South/Wales	0.170	-0.066	0.191	0.140	0.057	0.230	0.238	-0.061
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Scotland	0.376	-0.121	0.116	-0.121	0.022	0.483	0.301	0.037
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
London	0.219	-0.051	-0.187	-0.072	-0.008	-0.785	-0.362	-0.070
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Location mass point 1	0	0	0	0	0	0	0	0
Location mass point 2	0.834	0.423	-1.020	-0.973	-0.286	0.812	-0.755	-0.804
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Location mass point 3	-0.839	-0.333	0.272	0.005	0.835	0.986	-0.178	0.991
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Location mass point 4	-0.005	0.090	-0.748	-0.968	0.549	1.798	-0.933	0.187
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Value of the log-likelihood	-51,285.853							

Table 6a: Logit for probability of experiencing an accident – With Education

	Parameter estimate	Standard error
Intercept	0.065	0.003
Being employed	0.357	0.005
Female	-1.030	0.003
Age (divided by 10)	-1.702	0.003
Age squared (divided by 100)	0.224	0.002
Parental socioeconomic status at birth		
missing	0.054	0.003
High	-0.060	0.006
Low	-0.046	0.005
Mother smoked at pregnancy	0.085	0.004
missing	0.218	0.003

Mother's age at birth		
age (divided by 10)	-0.459	0.003
age squared (divided by 100)	0.706	0.003
missing	-1.088	0.003
Height at age 23	1.226	0.004
missing	1.937	0.003
Low birth weight	0.016	0.003
missing	-0.224	0.003
Math score at age 7	0.135	0.003
missing	0.032	0.003
Bristol Social Adjustment Guide at age 7	0.836	0.003
missing	-0.069	0.003
Region of residence at birth		
missing	0.246	0.003
North	0.045	0.003
Midlands	0	
South & Wales	0.021	0.004
Scotland	-0.105	0.003
London & South-East	0.039	0.003
Education		
O-level	0.018	0.010
A-level	0.142	0.005
Degree	0.012	0.005
Unobserved heterogeneity (factor loading)		
Probability 1: θ_1 θ_2	0.138	0.0004
Probability 2: (1-θ_1)θ_2	0.089	0.0002
Probability 3: θ_1 (1-θ_2)	0.471	0.0013
Probability 4: (1-θ_1)(1-θ_2)	0.302	0.0008
Location mass point 1	0	
Location mass point 2	1.246	0.003
Location mass point 3	-1.225	0.003
Location mass point 4	0.309	0.004

Table 6b: Multinomial logit with unobserved heterogeneity on transitions between work and disability states - With Education

	WD to NWD	NWD to WD	WND to WD	WND to NWD	WND to NWND	NWND to WD	NWND to NWD	NWND to WND
Intercept	-2.257	-2.175	-6.795	-7.177	-3.574	-3.157	-3.725	3.060
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Accidents	-0.137	0.153	0.818	1.454	0.029	0.845	0.852	0.144
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Female	0.788	-0.506	0.312	0.886	0.990	-1.191	-0.658	-0.885
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age (divided by 10)	-0.156	0.602	-0.089	-1.463	3.090	-0.779	-1.036	-2.450
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Age squared (divided by 100)	-0.034	-0.170	0.108	0.324	-0.669	0.101	0.245	0.360

	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.010)	(0.004)	(0.002)
Parental socioeconomic status at birth								
missing	0.192	-0.085	0.198	0.241	0.238	0.144	-0.508	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)
high	-0.090	-0.230	-0.095	-0.492	-0.045	0.390	-0.167	0.057
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
low	0.112	-0.059	0.138	0.215	0.257	0.505	-0.164	-0.092
	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)	(0.003)	(0.003)	(0.003)
Mother smoking during pregnancy	0.115	0.015	0.159	0.379	0.113	-0.194	0.175	0.018
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
missing	0.707	-0.469	0.142	0.550	0.229	-0.328	-0.386	-0.195
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)
Mother's age at birth								
age (divided by 10)	0.102	-0.398	-0.483	0.231	-0.331	0.064	-0.118	-0.065
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
age squared (divided by 100)	0.141	0.368	0.782	-0.583	0.528	0.110	0.253	0.200
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)
missing	-0.108	0.019	-0.344	-0.587	-0.412	-0.002	0.133	-0.201
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Height at 23	0.132	0.597	0.651	0.519	-1.251	0.055	0.570	0.747
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
missing	-0.396	0.302	-0.515	0.318	-2.020	0.113	-0.417	1.145
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LBW	0.017	-0.493	0.194	-0.203	-0.103	0.162	-0.182	-0.071
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
missing	0.153	-0.308	-0.388	-0.392	-0.216	0.564	0.100	-0.046
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Math score at age 7	-0.113	0.062	-0.069	-0.035	-0.543	0.002	-0.035	0.390
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
missing	0.031	-0.693	0.201	-0.009	0.036	-0.477	0.360	0.165
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
BSAG at age 7	0.509	-0.453	0.157	0.069	3.477	-0.053	0.111	-2.199
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
missing	-0.160	0.411	-0.061	-0.364	0.062	-0.062	-0.293	-0.136
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Region of residence at birth								
Missing	-0.017	-0.037	-0.172	-0.364	-0.218	0.027	0.106	0.507
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
North	0.387	-0.106	-0.064	0.180	0.188	0.120	0.393	-0.018
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
South/Wales	0.151	-0.153	0.165	0.094	0.017	0.169	0.225	-0.027
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Scotland	0.229	0.072	0.007	-0.315	0.203	0.433	0.349	-0.066
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)
London	0.064	-0.102	-0.134	-0.315	0.003	-0.400	0.328	-0.028
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Education								
O-level	-0.285	0.803	-0.245	-0.090	-0.500	0.129	-0.276	0.417

	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
A-level	-0.783	0.601	-0.291	-0.481	-0.747	-0.143	0.003	0.514
	0.003	(0.004)	(0.003)	(0.004)	(0.008)	(0.003)	(0.003)	(0.005)
Degree	-0.814	0.787	-0.596	-0.439	-0.878	0.035	0.134	0.737
	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)	(0.005)
Location mass point 1	0	0	0	0	0	0	0	0
Location mass point 2	-0.495	1.100	0.286	-0.920	-0.791	-0.938	-0.132	-0.095
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Location mass point 3	-1.489	0.356	-0.591	-1.694	-1.313	-0.401	-0.664	-0.376
	(0.003)	(0.004)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)
Location mass point 4	-1.984	1.456	-0.305	-2.614	-2.105	-1.339	-0.796	-0.472
	(0.005)	(0.005)	(0.005)	(0.004)	(0.008)	(0.005)	(0.004)	(0.006)
Value of the log-likelihood	-104,857.346							

Table 8: Multinomial logit on the initial state			
	Work/Disabled	Non-work/Disabled	Non-work/Non-disabled
Intercept	0.335	14.409	-0.457
	(1.843)	(3.636)	(1.128)
Gender	-0.351	-1.351	0.000
	(0.157)	(0.307)	(0.096)
Parental socioeconomic status at birth			
missing	0.048	-0.260	0.134
	(0.479)	(1.041)	(0.301)
high	0.107	-0.129	0.376
	(0.136)	(0.303)	(0.080)
low	0.193	-0.343	0.209
	(0.130)	(0.280)	(0.084)
Mother's smoking at pregnancy	0.135	-0.013	0.154
	(0.114)	(0.237)	(0.070)
Missing	-0.671	-0.170	0.213
	(0.648)	(1.030)	(0.276)
Mother's age at birth			
age (divided by 10)	0.131	-0.713	-0.312
	(0.737)	(1.594)	(0.493)
age squared (divided by 100)	-0.066	1.920	0.895
	(1.263)	(2.630)	(0.834)
missing	-21.074	-8.129	-10.135
	(3.047)	(1.755)	(1.102)
Height at 23	-1.962	-9.827	-1.429
	(0.792)	(1.554)	(0.481)
missing	-2.482	-15.868	-2.396
	(1.409)	(2.714)	(0.902)
LBW	0.396	0.550	0.172
	(0.205)	(0.346)	(0.141)
missing	-13.735	1.504	-0.236
	(1.764)	(1.003)	(0.594)
Math score at age 7	-89.830	-284.888	-13.059
	(25.459)	(59.619)	(15.389)
missing	-0.187	-2.264	0.384
	(0.465)	(0.704)	(0.273)
BSAG at age 7	16.479	36.636	20.903
	(6.220)	(12.157)	(4.040)
missing	-0.169	1.922	-0.021

	(0.462)	(0.663)	(0.264)
Region of residence at birth			
missing	35.601	6.170	10.698
	(3.610)	(1.909)	(1.115)
North	-0.159	0.418	0.344
	(0.143)	(0.322)	(0.092)
South/Wales	-0.108	0.731	0.273
	(0.161)	(0.343)	(0.105)
Scotland	-0.261	-0.042	0.195
	(0.196)	(0.448)	(0.122)
London	-0.373	-0.268	0.035
	(0.173)	(0.430)	(0.109)

Table 9: Disability rates per age for the different scenarios of accidents

	Scenario 1: Predicted by the model	Scenario 2: No accident	Scenario 3: Accident at 25	Scenario 4: Yearly accidents
Age 24	5.0%	4.8%	4.8%	6.9%
Age 25	5.1%	4.9%	5.3%	7.3%
Age 40	11.8%	11.0%	11.4%	21.0%

Table 10: Employment rates per age for the different disability scenarios

	Scenario 1: Predicted by the model	Scenario 2: No disability	Scenario 3: Disability at 25
Age 24	80.6%	81.1%	81.1%
Age 25	80.2%	80.7%	78.2%
Age 40	88.1%	90.5%	70.7%

Table 11: Disability rates per age for the different childhood scenarios

	Scenario 1: Predicted by the model	Scenario 2: Initial State	Scenario 2: Transitions	Scenario 3: Accidents
Age 20	4.2%	4.0%	4.0%	4.2%
Age 30	6.9%	6.7%	6.1%	6.8%
Age 40	11.8%	11.7%	10.0%	11.8%

Table 12: Employment rates per age for the different childhood scenarios

	Scenario 1: Predicted by the model	Scenario 2: Initial State	Scenario 2: Transitions	Scenario 3: Accidents
Age 20	84.3%	84.1%	87.3%	84.3%
Age 30	79.8%	79.8%	84.1%	79.8%
Age 40	88.1%	88.1%	91.4%%	88.1%

Appendix A

Labor Force Status

The labor force histories available in the NCDS are use to construct participants a measure of the labor force status at the beginning of each year. Since the survey participants were all born in March, we use March as the starting moment. The Centre for Longitudinal Studies (CLS) has transformed the data for waves 4 and 5 to include the detail of the economic activity for each month since the age of 16. In wave 6, we only have the starting dates and the economic status. We use this information to construct a monthly labor force status. The labor force status is divided between work and non-work spells. A work spell includes full and part-time employees and self-employed, voluntary work and maternity leave. It also includes apprenticeship schemes which are part of a job. Non-work spells include temporary and permanent sickness, prison time, traveling, retirement, and housework, government training schemes, unemployment, full and part-time education (as long as they are not in simultaneous employment) and traveling time.

We merge all the monthly information for all waves in order to fill in missing gaps. Nevertheless, for some participants missing data remains, especially because participants are not present in all subsequent waves. If the gap is more than a year, then the spell prior to the gap is treated as censored, and the data following the gap are not used in the estimation. Individuals must be present in wave 4, even if in the subsequent waves information is available about their entire labor history, because we need to control for their accident history since the end of age 16. For most individuals we will then have information since the age of 16 until they are censored because of attrition, missing data or the last interview. Finally, we exclude the time while finishing education and start the record since their first job. For both accidents and disability (and hospitalizations), the data includes information on the timing of the event and this is matched to the corresponding work or non-work spell. Because the information for disability and accidents is recorded yearly, our final dataset contains the yearly records of labor force, disability and accidents.

Disability

We base our definition of disability on the Handbook of Health Economics as the mental and physical characteristics that, either constrain normal daily activities, or cause a substantial reduction in productivity on the job. The NCDS data contains a set of question on health status. Individuals are asked at ages 23, 33 and 42 whether they have a longstanding illness, disability or infirmity which limits their activities compared to people their own age. They are subsequently requested to document whether it limits their daily activities or the work they can do, the age of the disability onset and the type of disability. Disability types are coded according to the international classification of disease (ICD) produced by the World Health Organization (1977).

The ICD is extensively used in health studies and is grouped into 17 broad categories:

1. Infections and parasitic diseases (e.g. tuberculosis, shingles, herpes simplex, glandular fever),
2. neoplasms (e.g. Hodgkin's disease, leukemia),
3. endocrine, nutritional and metabolic diseases and immunity disorders (e.g. obesity, diabetes),
4. diseases of the blood and blood-forming organs (e.g. anemia, coagulation defects),
5. mental disorders (e.g. depression, neurotic disorders, mental retardation),
6. diseases of the nervous system and sense organs (e.g. epilepsy, migraine, blindness, deafness),
7. diseases of the circulatory system (e.g. hypertension, pericarditis, aortic aneurysm),
8. diseases of the respiratory system (e.g. bronchitis, asthma, pleurisy),
9. diseases of the digestive system (e.g. duodenal ulcer, appendicitis, cirrhosis of the liver),
10. diseases of the genitourinary system (e.g. renal failure, cystitis, infertility),
11. complications of pregnancy, childbirth and the puerperium (e.g. spontaneous abortion, etopic pregnancy),
12. diseases of the skin and subcutaneous tissue (e.g. eczema, psoriasis),
13. diseases of the musculoskeletal system and connective tissue (e.g. rheumatoid arthritis, derangement of joint)
14. congenital anomalies,
15. certain conditions originating in the Perinatal period,
16. symptoms, signs and ill-defined conditions,
17. Injury and poisoning (e.g. fractures, sprains, dislocations, traumatic amputation).

Education

The cohort students followed an education system where they were required to pass an exam at age 11 which determined their educational path. If they succeeded, they would go to a grammar school and follow a university track. They and prepare for public examination in different subjects: ordinary "O-level" exams at age 16 and advanced "A-levels" at age 18. Students are admitted to universities based on their performance at A-level exams. If they could not enter grammar schools they would go to secondary schools an obtain certificate of secondary education (CSE), after which they can enter the labor market. General vocational qualifications are also available and have equivalence to the "O-levels" and "A-levels".

Appendix B

Table A1: Test on non-random attrition: Logit of participation in wave 5 on health and labor market status in wave 4

Variables	Coefficients	Z-values
Employed at age 23	0.616	(13.65)
Disabled at age 23	0.265	(2.73)
Female	0.277	(4.54)
Parental socioeconomic status at birth		
Missing	0.227	(1.17)
High	0.139	(2.66)
Low	-0.184	(3.57)
Mother smoked after the fourth month of pregnancy		
Missing	-0.309	(1.73)
Yes	-0.078	(1.76)
Mother's age at birth (in years)	0.561	(1.81)
Missing	0.450	(0.45)
Mother's age squared at birth (in years)	-0.922	(1.73)
Height at age 23 (in meters)	0.820	(2.73)
Missing	0.699	(1.25)
Birth weight		
Missing	-0.279	(0.92)
Low (less than 2500 grams)	-0.193	(2.12)
Math test score at age 7 (scale 0-10)	56.460	(8.18)
Bristol Social Adjustment Guide at age 7	-15.906	(6.93)
Region of residence at birth		
Missing	0.325	(0.33)
North	-0.015	(0.27)
South & Wales	0.102	(1.55)
Scotland	-0.141	(1.91)
London & South-East	-0.020	(0.31)
Constant	-2.224	(3.09)
Observations	12448	

Table A2: Test on non-random attrition: Logit of participation in wave 6 on health and labor market status in wave 4

Variables	Coefficients	Z-values
Employed at age 23	0.507	(11.78)
Disabled at age 23	0.230	(2.58)
Female	0.204	(3.61)

Parental socioeconomic status at birth		
Missing	-0.0200	(0.12)
High	0.171	(3.57)
Low	-0.148	(3.05)
Mother smoked after the fourth month of pregnancy		
Missing	-0.250	(1.46)
Yes	-0.089	(2.13)
Mother's age at birth (in years)	0.450	(1.55)
Missing	0.867	(0.88)
Mother's age squared at birth (in years)	-0.718	(1.44)
Height at age 23 (in meters)	0.800	(2.86)
Missing	0.569	(1.08)
Birth weight		
Missing	-0.307	(1.07)
Low (less than 2500 grams)	-0.147	(1.69)
Math test score at age 7 (scale 0-10)	64.046	(10.02)
Bristol Social Adjustment Guide at age 7	-18.428	(8.38)
Region of residence at birth		
Missing	0.029	(0.03)
North	-0.058	(1.11)
South & Wales	0.102	(1.68)
Scotland	-0.076	(1.10)
London & South-East	-0.037	(0.62)
Constant	-2.381	(3.54
Observations	12448	