The interrelation between disability and work and the role of health shocks

Maarten Lindeboom[a]

Ana Llena-Nozal[b]

Bas van der Klaauw[c]

April 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 the unanticipated nature of accidents to estimate the causal effects of health shocks. 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 an accident. And the occurrence of accidents increases 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 or of getting out of work.

[a] Free University Amsterdam & Tinbergen Institute.
[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.
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 (e.g. Smith 1998, 1999). Better-educated, high-income people generally have a better health and lower mortality rates. As a key element to the association between health and socioeconomic status early childhood conditions or health shocks are often mentioned (e.g. Currie and Hyson, 1999). However, a large range of the literature is based on reduced-form studies that offer little consensus about the underlying mechanisms (see for example the discussion in Case, Fertig and Paxson, 2005).

During adulthood health deteriorates with age and the rate of deterioration is influenced by decisions made in the course of the life cycle and by shocks. Labor market choices are important because they affect health directly and indirectly. Adverse working conditions affect the rate at which health depreciates. Employment may also trigger negative health shocks, such as accidents and these shocks may lead to a disability that restrict individuals in doing their daily and/or work activities. This may in turn affect the individual's labor supply decisions and work outcomes.

Smith (1999) describes the ongoing debate about the direction of the causal relations between health and socioeconomic status. In general it is very difficult to disentangle the underlying causal mechanisms, largely because unobservables may also relate to health and work outcomes and because we lack suitable instruments (e.g. Smith, 1999). This stresses not only the difficulties that arise in identifying causal relations, but also the importance of actually disentangling the causal relation between socioeconomic status and health. Identification requires independent variation in either health or work outcomes to assess the effect of one on the other; see for example Lindahl (2005) who uses lottery prize winning to study the effect of income on health.

In this paper we focus on the interrelation between disability and work, and the role of health shocks. We use accidents as a measure of health shocks and exploit the unanticipated nature of accidents to identify some of the causal mechanisms. In particular, we investigate the consequences of health shocks on labor market outcomes and the onset of disabilities. We define a disability as a permanent chronic condition that restricts individuals in their daily activities and/or in their work. Currently in the UK there are 7.1 million individuals (3.7 million men and 3.4 million women) with a disability and only about half of disabled people of working age are in employment (Smith & Twomey, 2002). Re-employment probabilities are very low for this group.

Our sample follows workers up to age 40 and at that age already 12.63% of the workers face one or more disabilities. About 17% of the disabled are out of work. Bad health may be an important factor for the work decision when people are in their prime age, but even more so when they reach the age that they become eligible for (early) retirement programs (see the surveys of Lumsdaine & Mitchell, 1999, Bound & Burkhauser, 1999). Disabilities that cause inactivity of at later ages often show up at earlier ages.

In the empirical analyses we use an event history model for the interrelation between work, disability and accidents. Individuals experience accidents at different moments in life and beforehand they are not aware that an accident will take place. To identify the causal effect of accidents on the work and disability status, we require that there is unanticipated variation in the timing of accidents. It should be stressed that we do not require accidents to be exogenous (conditional on a set of observed characteristics), but instead we allow for unobservables that simultaneously affect employment probabilities, the onset of disability and the likelihood of having an accident. See Abbring and Van den Berg (2003) for an extensive discussion on the identification of treatment effects in dynamic models. Unanticipation in this context means that in advance the exact timing of an accident is unknown to the individual. It does not rule out that individuals may be aware that at some moments the risk of having an accident is higher than in other periods or that the risk of accidents differs between individuals. In particular, a substantial share of the accidents is related to work, which we take into account in our model framework by allowing the probability of experiencing an accident to depend on the current employment status.

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. These individuals are followed from birth up to 2000. This is only one of the few data sets that track individuals from birth until middle age. The NCDS contains abundant information on the situation of the family where the individual was born in and early childhood health outcomes.

To illustrate the mechanisms and the importance of health shocks, we perform some simulation experiments using our estimated model. In these simulation experiments 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 to what extent the relation between childhood circumstances and health and labor market outcomes during adulthood is driven by health shocks.

The structure of the paper is as follows. Section 2 presents some theory and the estimation strategy. 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.

2. Theoretical background and empirical model specification

2.1 Theoretical Background

In this section we present the statistical model, which we use in the empirical analyses. Before giving the outline of the model, we first discuss some hypotheses that may explain the interrelation between early childhood characteristics, health shocks, labor market outcomes and the onset of disability.

Smith (1998, 1999) shows that there exists a strong association between health and wealth at adulthood. This health-wealth gradient holds within all age groups, and in particular those with low health experience lower wealth growth over the life cycle. Case, Fertig and Paxson (2005) give a simple theoretical model that provides some insight in the possible mechanisms underlying the association between health and socioeconomic outcomes during adulthood. Poor childhood health is often considered to be an important contributor to association between health and economic outcomes (e.g. Case, Lubotski and Paxson, 2002, Currie and Hyson, 1999, Currie and Stabile, 2003, Dobblhammer, 2003). Early childhood health can influence later life morbidity and mortality because adverse conditions during childhood may increase disease risks in later life (Barker 1995). Also in the model of Case, Fertig and Paxson (2005) early childhood circumstances are a key factor.

Recent literature focused on the importance of early childhood health and socioeconomic conditions on later health and labor market outcomes. Currie and Hyson (1999) investigate the consequences of low birth weight. They mention that children from a low socioeconomic status both suffer from more often from a low birth weight and are less likely to recover from it. Case, Fertig and Paxson (2005) find that controlling for parental income, education and social class, children with poorer uterine environments and poorer childhood health have significant lower educational attainment, poorer health and lower socioeconomic status as adult. It is well known that individuals from lower socioeconomic backgrounds are more often involved in unhealthy behavior such as smoking and obesity, but Smith (1999) argues that behavioral differences between individuals from different socioeconomic background cannot fully explain the differences in health outcomes. Neither can differences in access to health care explain the differences in health outcomes.

Individuals from adverse early childhood circumstances often have lower educational attainments, and the level of education is correlated with health during adulthood (Fuchs, 2004). Marmot et al (2001) argue that adverse conditions during childhood may negatively influence teenage health and health and socioeconomic status during early adulthood and therefore later life outcomes. They particularly stress that adult socioeconomic status is the most important factor in explaining health during adulthood. Childhood circumstances and health at early childhood are generally not as important. This implies that education, health and labor market outcomes at early adulthood are intermediate variables for early childhood indicators. Case, Fertig and Paxson (2005) find indeed that parental socioeconomic status and early childhood health influence health and socioeconomic outcomes mainly through early adult outcomes. But controlling for education, health and socioeconomic status at early adulthood, chronical conditions in childhood and the prenatal environment appear to have a significant effect on health at middle age.

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). Individuals with poorer childhood health and from a lower socioeconomic background may more often be exposed to stress, which affects health in the long-run. Health could also be considered as an element in the human capital function. Health directly affects labor market outcomes and thus the rate at which skills are accumulated. Individuals with poor initial health therefore experience a slower rate of wage growth and are more likely to be non-employed. 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 negative health shocks.

The health demand model developed by Grossman (1972, 1999) assumes that individuals 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 behavior and past consumption. Individuals include health in their maximization, which implies that they must have expectations about their health trajectories based on information on their previous health, their age, their health behavior and their genetic component. As new health events occur, people update their expectations about their health trajectories and change behavior accordingly. This stresses immediately the difficulties in identifying the causal relations between health and labor market outcomes. If health depreciation or health shocks are anticipated, then individuals have already changed behavior. However, the health path is not entirely predictable, the timing of the occurrence of health shocks is often unknown in advance.

Smith (1998) stresses the importance of shocks in disentangling the causal relation between health and socioeconomic status. Shocks should contain new information to an individual and thereby provide some exogenous variation in either health or socioeconomic status. What is important is the unanticipated nature of a shock: people may anticipate some onset, but the timing of the actual realization of the shock should be unanticipated. Economic literature using natural experiments mainly focused on the effect of socioeconomic status on health outcomes. For example, Lindahl (2005) uses lottery prizewinners to investigate the effect of wealth on health. His empirical results show that wealth has a significant effect on health.

There is less literature that exploits exogenous variation in health to investigate the causal relation from health to socioeconomic status. Smith (1998) suggests using the onset of chronical conditions and hospitalization as unanticipated health shocks. Adams, Hurd, McFadden, Merrill and Ribeiro (2003) conclude that these types of health shocks only marginally affect socioeconomic outcomes. However, their sample of elderly only allows for very limited effects on labor market outcomes. Using hospitalization and the onset of diseases might be somewhat problematic in a US setting. Smith (1998) mentions that only half of the individuals are fully insured. Non-insured individuals have to pay for medical care and therefore the choice to go to hospital might be related to the individuals' financial situation. In particular, wealthy people might go to hospital earlier than poor people. And hospitalization for non-insured individuals has a direct negative effect on wealth, which does not go via health depreciation. This suggests heterogeneity in the effect of hospitalization on health. Indeed Smith (1999) shows that the impact of a new health onset is larger on individuals without health insurance than with health insurance.

In our model framework we use unanticipated accidents to investigate the effect of these health shocks on labor market status and disability status. 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. Since health shocks occur at different moments in life, the model should be dynamic. And the discussion above suggests that we should acknowledge the importance of early childhood conditions. 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 health and labor market outcomes at early adulthood, which again influence health and labor market outcomes at later ages. If the hypothesis of Marmot et al (2001) would be true, this would be the only relevant effect of early childhood conditions. Second, early childhood conditions can directly affect the rate of health depreciation and the incidences and length of workless spells. This would be consistent with the idea to consider health as input in the human capital function. Finally, the probability of experiencing an accident during the course of life is allowed to depend on early childhood conditions. This implies that adverse childhood conditions may be trigger for later health shocks, which in turn influence health and labor market outcomes during adulthood.

A dynamic model has the advantage that we can relax the requirements for accidents to be valid health shocks substantially. Smith (1998) suggests that all risk factors of experiencing accidents should be included. Within our dynamic model we allow accidents to be endogenous, i.e. we allow for unobserved heterogeneity that affects both the probability of having an accident and disability and labor market outcomes. The advantage of a dynamic model is that if accidents are unanticipated, the effect of an accident can be identified without exclusion restrictions or strong functional form restrictions (e.g. Abbring and Van den Berg, 2003).

2.2 Empirical specification

In this section we describe our empirical model, which we use to evaluate the effects of health shocks on the onset of disabilities and labor market status. Before discussing our model, we first provide a brief outline of the data, the next section gives a more detailed description of the data. From the database we have constructed 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 on 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). Individuals can in each period move between working and non-working. 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 depends on the individual's current labor market state, accidents can be work related and therefore employed individuals might have higher probabilities of getting an accident. The probability that an accident occurred 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 $(i,j)\neq(k,m)$ 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. 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 $(w_1,w_2,\ldots,w_N)$, in which each element w_n has two discrete mass points at 0 and 1. 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 $s_l(1), s_l(2),\ldots, s_l(T)$ and the health states of the individual are given by $s_h(1), s_h(2),\ldots, s_h(T)$ and the sequence $a(1),a(2),\ldots,a(T)$ 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 might know that in particular periods the probability of getting an accident is high, for example when they are working. 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. 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 only consider 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[1]. After selecting only those with complete labor and health histories, our final sample consists of 12,448 individuals. Case, Fertig and Paxson (2005) investigate 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. 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 (see Kerkhofs Lindeboom, 1997 for a discussion of this test). We also included the set of individual characteristics as controls. We performed the same test for attrition from the 1999/2000-wave. The *p*-values for joint significance of the labor market and health status in 1981 on attrition in 1991 and 1999/2000 are 0.000 and 0.000 respectively. The attrition is particularly selective with respect to labor market status[2].

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 5, when almost everyone has left full-time education. The fraction of males out of the labor force slowly increases with age. The unemployment rate is relatively constant except for the ages 22 until 24,

[1] 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)

[2] See Tables A1 and A2 in the Appendix B

when there seems to be some increased unemployment. This might either be related to business cycle effect, i.e. the recession in the late 1970s/beginning 1980s or to 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 forces starts to decrease, employment rates increase. The key difference between the patterns of men and women is thus the fraction being out of the labor force.

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 (WHO, 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 Department for Work and Pensions (UK) 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.

We use accidents as unanticipated health shocks. An accident is categorized as unanticipated event after which an individual is admitted to hospital or attending a hospital outpatient or casualty department. 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.[3] Men are much more likely to experience 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 type of accidents does not coincide. The fact that a large share of the accidents is workplace accidents shows the importance to control in the empirical model for the labor market status of the individual in specifying the probability of having an accident. 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 is very rich on individual characteristics. For each individual we observe a range of variables that give information on an individual's initial health assets, the socioeconomic status during early childhood and cognitive ability at 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 on these variables. For many variables there is some item non-response. To avoid losing many observations we follow Case, Fertig and Paxson (2005) by constructing dummy variables that indicate if the information on a variable is missing.

---

[3] 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.

Infants with a birth weight below 2500 grams are considered to have low birth weight, for which we construct a dummy variable. According to the World Health Organization (2004), epidemiological research shows that babies with low birth weight are more likely to die than heavier babies. Low weight at birth is the result of either preterm birth (before 37 weeks of gestation) or restricted growth. Short gestation is the main cause of death, morbidity and disability. Low birth weight due to restricted growth is closely associated chronic diseases and cognitive development problems. Next we consider the height at age 23, which is an indicator for poor living conditions during childhood. 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. Therefore, we also include mother's age squared in the empirical analyses.

The parental socioeconomic status describes 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.[4] Higher scores indicate higher maladjustment. The data also included information on the

[4] 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.

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 than O-levels, O-level equivalent, A-level equivalent, and degree equivalent (see Case, Ferig and Paxson, 2005). We return to the issue whether health affects educational outcome or the other way around in the next section. Finally, we will use the region at birth to control for differential factors in geography, industry, living conditions, etc.

4. Empirical results

4.1 *Parameter estimates*

In this section we discuss our estimation results. First, we present the parameter estimates of a model specification that does not include education as explanatory variable. Next we do include education as explanatory variable. Finally, we estimate the model separately for men and women.

The parameter estimates are reported in Tables 4a and 4b. The unobserved heterogeneity is significant and the preferred specification is a factor-loading specification with two elements that each take two values, i.e. the vector w of random variables specified in Subsection 2.2 has two elements (w_1,w_2). Within each transition probability there are four mass point, which are due to the factor loading specification somewhat related to each other. Most probability mass is located at a mass point (location 3) describing individuals with a low probability of experiencing an accident. Individuals who are most likely to get an accident (mass point location 2) are also more likely to switch states than the majority of the individuals. Both other mass points describe individuals who have an average probability of experiencing an accident, but are either not very likely to switch labor market and disability status (mass point location 4) or are much more likely than other individuals to changes states (mass point locations 1).

Table 4a shows the logit specification for the probability of experiencing an accident. Being employed raises the probability of suffering an accident. Recall from Table 1 that indeed a substantial share of the accidents is workplace related. Males are much more likely to get an accident than females. Recall from Subsection 2.2 that men experienced more accidents. Obviously the differences in employment rates between men and women and the differences in observed individual characteristics cannot explain the differences in accident incidences between men and women. The probability of having an accident is U-shaped in age; it is decreasing until age 38 and increasing afterwards.

Both the variables describing the individual health at birth and the cognitive ability are important in explaining differences in the likelihood of getting an accident. 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. This is consistent with the life cycle model, which states that the rate of health depreciation differs between individuals with different early childhood conditions, because negative health shocks during childhood are a trigger for 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 also have higher probabilities of getting an accident. It is difficult to connect strong causal interpretations to these effects as for 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 the multinomial logit specifications for the transition probabilities between the different labor market and health states. The key parameters of interest are the parameters describing the effects of having an accident on the transition probabilities. Accidents have a significant impact on all transitions probabilities and in most cases having an accident increases the mobility between different states. The effects of accidents are most substantial on the probability of a transition from the non-disability states into the disability states. We do not find any unambiguous effect on how accidents affect changes between labor market states. In Section 5 we show some simulation experiments to investigate the magnitude of the effect of accidents on labor market and health outcomes.

Being female increases the transition rate from the work states towards the non-work states and decreases the transition rates in the opposite direction. The reason women have lower employment rates is thus not only that women start their careers more often in a non-work state, but also that if they are working, they are more likely to quit working. Furthermore, when women are working, they are more likely to become disabled than men. Non-working women have lower probabilities to become disabled than non-working men. 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. And when non-employed, these individuals have on average short non-employment spells. Relatively tall people at age 23 are more likely to become disabled than shorter people, but while not being disabled 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 estimate our model separately for men and women. Tables 5a and 5b report the estimation results for respectively men and women.
(SEPARATELY ESTIMATION FOR MEN AND WOMEN)

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 as regressor. There are no important changes in the magnitude and significance of the other variables once education is included. On the other hand, the coefficients for the education dummies are significant. The control 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 as an additional individual characteristic 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.

(ESTIMATION RESULTS WITH ATTRITION DUMMY)

5. Simulation experiments

5.1 *Initial state*

In this section we perform some simulation experiments to investigate the importance of health shocks on labor market outcomes 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. For that reason, age did not have a significant impact on the initial state. 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 is in agreement with Currie and Hyson (1999), who only find modest effects of early childhood conditions on health and labor market outcomes at early adulthood. It indicates that the fetal origins hypothesis (see Marmott et al, 1999), that the effects of early childhood conditions on middle age outcomes mainly works via the socioeconomic status at early childhood, is not the most important explanation for the relation between early childhood conditions and health and socioeconomic outcomes during the life-cycle.

5.2 Simulations

From the estimates in Tables 8, 4a and 4b, we can derive the relationship between the work and disability states and the different regressors such as age, socioeconomic status and other childhood characteristics. We can therefore depict the evolution of the probabilities of being in each state over time. Figure 6 depicts work and disability probabilities over age, starting at age 16. The figure shows a much higher probability of being in the work nondisabled state. Nevertheless, this probability first declines as people age as there is first an increased chance of being in a NWND state. In the late 20s, the probability of being in a WND states increases again as the chances of nonwork decline, but this increase is mitigated by the rise in disability (the increase in the probability of WD and NWD). We can see that this prediction matches the observed probabilities depicted in Figure 5. Our model tends to smooth the evolution of probabilities but represents a satisfactory portray of reality.

We use the coefficients from the estimation to perform several experiments. Firstly, we investigate the effect of disability on employment and to this purpose we do the following: (a) simulate the model once with everybody disabled as initial state and second everybody not disabled as initial state. (b) no-one is disabled until age 25 and then everyone becomes disabled. Secondly, we look at the effect of accidents through different scenarios: (a) nobody gets an accident, (b) everybody gets an accident at age 25, (c) decrease accident incidence with 25%, (d) reduce workplace accidents (coefficient of work on accident = 0). Finally, we explore the role of childhood conditions and make everybody come from a high SES background (plus no low birth-weight and no maternal smoking) during the (a) initial state, (b) direct transition probabilities, (c) accident probabilities in order to compare where early childhood conditions are most important. The results are shown in Tables 9, 10 and 11 respectively.

In Table 9, we depict different disability scenarios. In the first case, everybody is disabled initially and there is thus no probability of moving back to a non-disabled state. We see from the table and from figure 7a that the probability of working while disabled decreases until it reaches a sort of plateau from the age of 22 onwards. In the case where people leave full-time education without disability, Figure 7b and table 9 show that the probability of becoming disabled rises steeply. By the age of 40 this probability is getting substantially close to the standard situation. The last simulation for disability is the case where everybody gets disabled at the age of 25. Figure 7c depicts this scenario. Until age 25 there is a situation of full-employment. Once people become disabled there is a dramatic decrease in the probability of being employed, decreasing to 0.0327 by the age of 40.This suggests that if people do not get ill health nor health shocks, there will be no situation of non-employment. We simulate this scenario in figure 9, where we observe that, indeed, this will be the case.

From Table 10 can be observed that reducing accidents greatly reduces the probability of being disabled and employed. We see that if everybody gets an accident at age 25, by the age of 40 the likelihood of being disabled is slightly higher than if people do not experience any accidents. This could be explained by the fact that our sample consisted of individuals who suffered from many accidents and that perhaps the likelihood of disability increases after repeated health shocks. Besides, reducing workplace accidents appears to be more important in reducing the likelihood of experiencing a disability than reducing the incidence of all accidents by 25%.

In Table 11, we depict the effect of childhood characteristics at different stages. We observe than in our three scenarios the probability of being disabled is lower than in the standard predictions. In all cases we observe a decrease in the likelihood of becoming disabled and an increase in the probability of being at work while not disabled. Childhood characteristics appear to matter more in the transition probabilities and have the least impact on the probability of experiencing accidents.

6. Conclusion

This paper explores the relationship of disability and work over the life cycle. We are particularly interested in the influence of childhood and early adulthood on later outcomes. In a dynamic model, we exploit the occurrence of unanticipated health shocks, such as accidents that affect both the disability status and labor market outcomes. Even though an accident is an unanticipated event to the individual, the incidence of accidents depends on the current labor market status and both observed and unobserved characteristics. Within an event history model, we can identify the causal effect of accidents on disability and labor market outcomes. To estimate the model, we use data from the British National Child Development Study (NCDS).

The results indicate that indeed current labor market status positively influences the probability of experiencing health shocks and that individuals who have more disadvantaged childhood and early adulthood characteristics are more likely to have accidents. Accidents have a strong impact on the individual's outcomes; in particular, the occurrence of a disability if greatly affected by accidents. Furthermore, individuals with disabilities have a much higher probability of entering unemployment. To some degree the higher probability of experiencing negative health shocks explains that individuals with less initial health assets are more likely to be unemployed and have disabilities.

Our results are particularly relevant for policy matters as we postulated in the introduction. We see that the impact of childhood and background characteristics is indeed twofold: early childhood factors influence both the rate of accidents and the probability of getting a disability, increasing thus the chances of being disabled directly and indirectly. This is partly in line with previous literature findings (Case, Lubotski & Paxson, 2001; Case, Fertig & Paxson, 2005; Currie & Hyson, 1999) where it has been found that health in childhood influences earnings and heath in adulthood via educational attainment. Nevertheless, our findings suggest that the impact is not on initial employment and disability status but into later outcomes but rather a trigger for accumulation of disadvantages during adulthood. This conclusion is important for public policy since it implies that a policy that improves early childhood outcomes for the economically disadvantages or for those who have experienced health shocks at early ages, will reduce the odds of experiencing adverse health shocks later in life as well as the probability of getting 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 estiamtes 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.

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, 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.

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	WD(t)	NWD(t)	WND(t)	NWND(t)	Female	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 accidents

	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 (mass points)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 6a: Logit for probability of experiencing an accidents – 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: $\theta_1\ \theta_2$	0.138	0.0004
Probability 2: $(1-\theta_1)\theta_2$	0.089	0.0002
Probability 3: $\theta_1\ (1-\theta_2)$	0.471	0.0013
Probability 4: $(1-\theta_1)(1-\theta_2)$	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 (mass points)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	0	0	0	0	0	0	0	0

point 1								
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: Calculations with the model- Disability

Scenario		Standard	All disabled on initial state	Nobody disabled on initial state	Disability at age 25
Predicted probabilities Age			Work Disabled		
	16	0.0325	0.8132	0.0000	0.0000
	20	0.0299	0.7394	0.0053	0.0000
	30	0.0471	0.7045	0.0258	0.0378
	40	0.0821	0.7065	0.0622	0.0327
Scenario		Standard	All disabled on initial state	Nobody disabled on initial state	Disability at age 25
Predicted probabilities Age			Non-Work Disabled		
	16	0.0122	0.1868	0.0000	0.0000
	20	0.0120	0.2605	0.0023	0.0000
	30	0.0215	0.2955	0.0124	0.0622
	40	0.0361	0.2934	0.0278	0.0673
Scenario		Standard	All disabled on initial state	Nobody disabled on initial state	Disability at age 25
Predicted probabilities Age			Work Non-Disabled		
	16	0.8683	0.0000	0.9083	0.9083
	20	0.8135	0.0000	0.8424	0.9938
	30	0.7511	0.0000	0.7754	0.0000
	40	0.7990	0.0000	0.8245	0.0000
Scenario		Standard	All disabled on initial state	Nobody disabled on initial state	Disability at age 25
Predicted probabilities Age			Non-Work Non-Disabled		
	16	0.0869	0.0000	0.0917	0.0917
	20	0.1446	0.0000	0.1499	0.0062
	30	0.1803	0.0000	0.1863	0.0000
	40	0.0827	0.0000	0.0854	0.0000

Table 10: Calculations with the model- Accidents

Scenario		Standard	Nobody gets an accident	everybody gets an accident at age 25	decrease accident incidence with 25%	reduce workplace accidents
Predicted probabilities Age				Work Disabled		
	16	0.0325	0.0325	0.0325	0.0325	0.0325
	20	0.0299	0.0290	0.0290	0.0297	0.0291
	30	0.0471	0.0441	0.0466	0.0463	0.0444
	40	0.0821	0.0763	0.0786	0.0806	0.0768
Scenario		Standard	Nobody gets an accident	everybody gets an accident at age 25	decrease accident incidence with 25%	Reduce workplace accidents
Predicted probabilities Age				Non-Work Disabled		
	16	0.0122	0.0122	0.0122	0.0122	0.0122
	20	0.0120	0.0117	0.0117	0.0120	0.0118
	30	0.0215	0.0204	0.0216	0.0212	0.0206
	40	0.0361	0.0341	0.0353	0.0356	0.0343
Scenario		Standard	Nobody gets an accident	everybody gets an accident at age 25	decrease accident incidence with 25%	Reduce workplace accidents
Predicted probabilities Age				Work Non-Disabled		
	16	0.8683	0.8683	0.8683	0.8683	0.8683
	20	0.8135	0.8145	0.8145	0.8138	0.8150
	30	0.7511	0.7542	0.7518	0.7519	0.7546
	40	0.7990	0.8061	0.8031	0.8008	0.8059
Scenario		Standard	Nobody gets an accident	everybody gets an accident at age 25	decrease accident incidence with 25%	reduce workplace accidents
Predicted probabilities Age				Non-Work Non-Disabled		
	16	0.0869	0.0869	0.0869	0.0869	0.0869
	20	0.1446	0.1447	0.1447	0.1446	0.1441
	30	0.1803	0.1812	0.1799	0.1805	0.1803
	40	0.0827	0.0835	0.0830	0.0829	0.0829

Table 11: Calculations with the model- Childhood characteristics

Scenario		Standard	High SES during initial state	High SES during transition probabilities	High SES during accident probabilities
Predicted probabilities Age				Work Disabled	
	16	0.0325	0.0299	0.0325	0.0325
	20	0.0299	0.0286	0.0300	0.0298
	30	0.0471	0.0462	0.0441	0.0469
	40	0.0821	0.0813	0.0735	0.0818
Scenario		Standard	High SES during initial state	High SES during transition probabilities	High SES during accident probabilities
Predicted probabilities Age				Non-Work Disabled	
	16	0.0122	0.0110	0.0122	0.0122
	20	0.0120	0.0114	0.0101	0.0120
	30	0.0215	0.0210	0.0166	0.0214
	40	0.0361	0.0356	0.0266	0.0360
Scenario		Standard	High SES during initial state	High SES during transition probabilities	High SES during accident probabilities
Predicted probabilities Age				Work Non-Disabled	
	16	0.8683	0.8600	0.8683	0.8683
	20	0.8135	0.8120	0.8431	0.8136
	30	0.7511	0.7519	0.7972	0.7512
	40	0.7990	0.8000	0.8407	0.7994
Scenario		Standard	High SES during initial state	High SES during transition probabilities	High SES during accident probabilities
Predicted probabilities Age				Non-Work Non-Disabled	
	16	0.0869	0.0990	0.0869	0.0869
	20	0.1446	0.1480	0.1167	0.1445
	30	0.1803	0.1809	0.1420	0.1804
	40	0.0827	0.0829	0.0591	0.0828

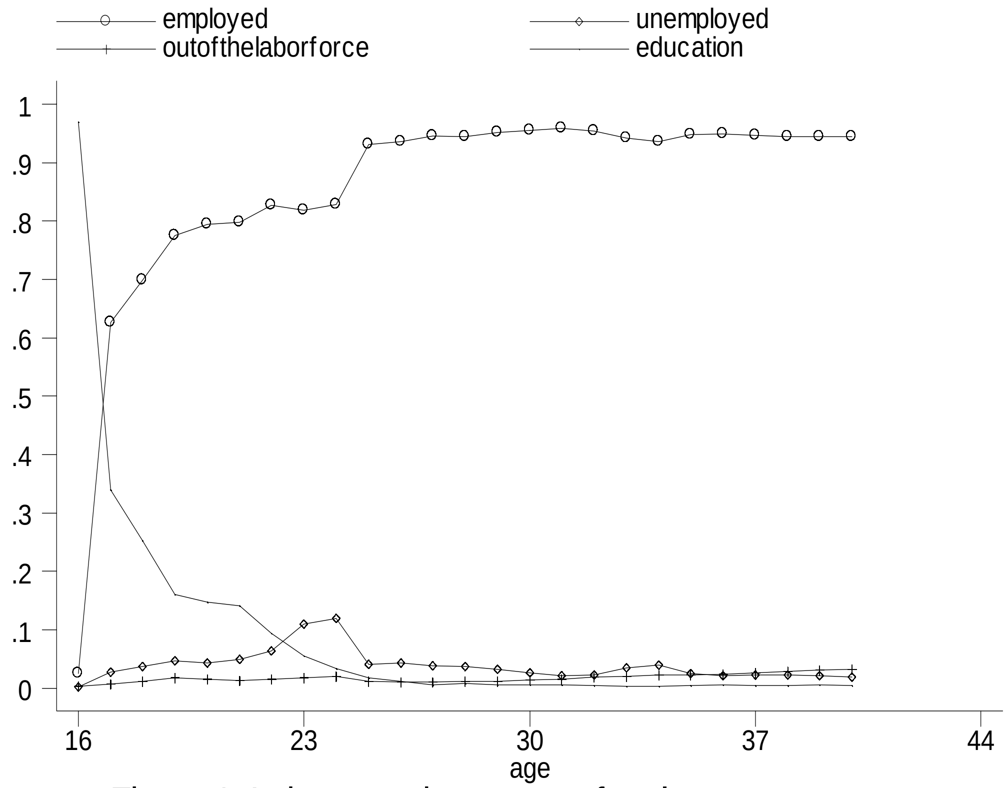


Figure 1: Labour market status of males per age

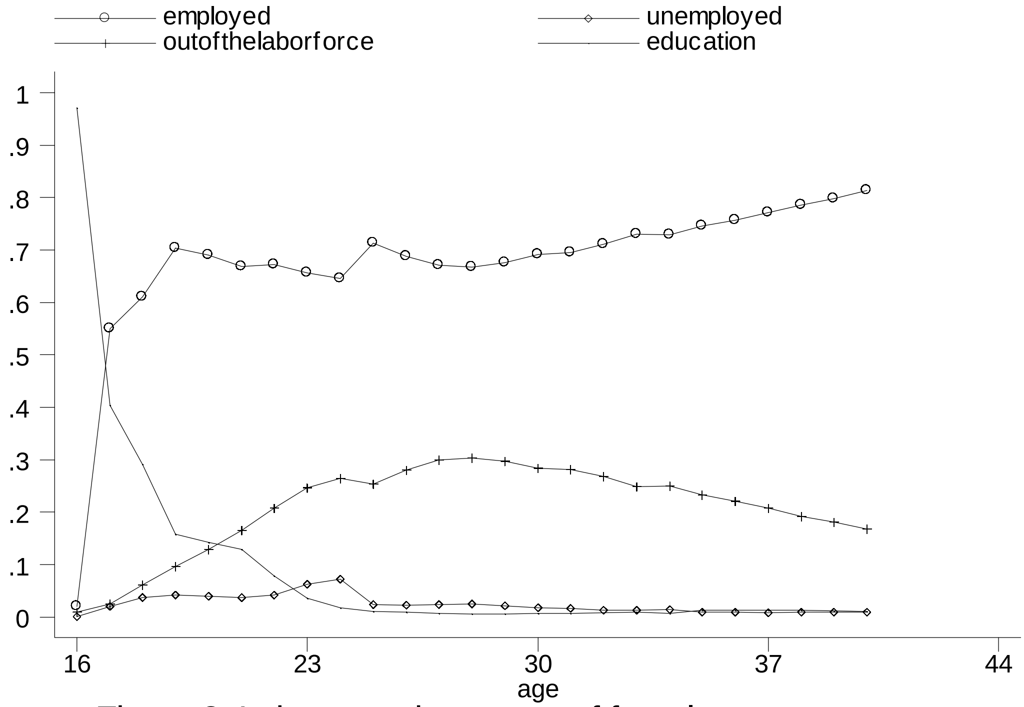


Figure 2: Labour market status of females per age

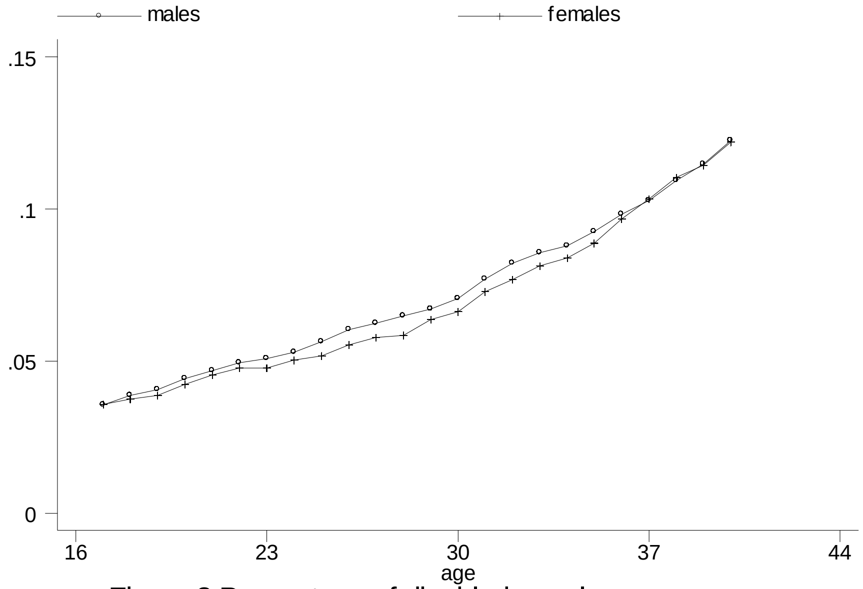


Figure 3:Percentage of disabled people per age

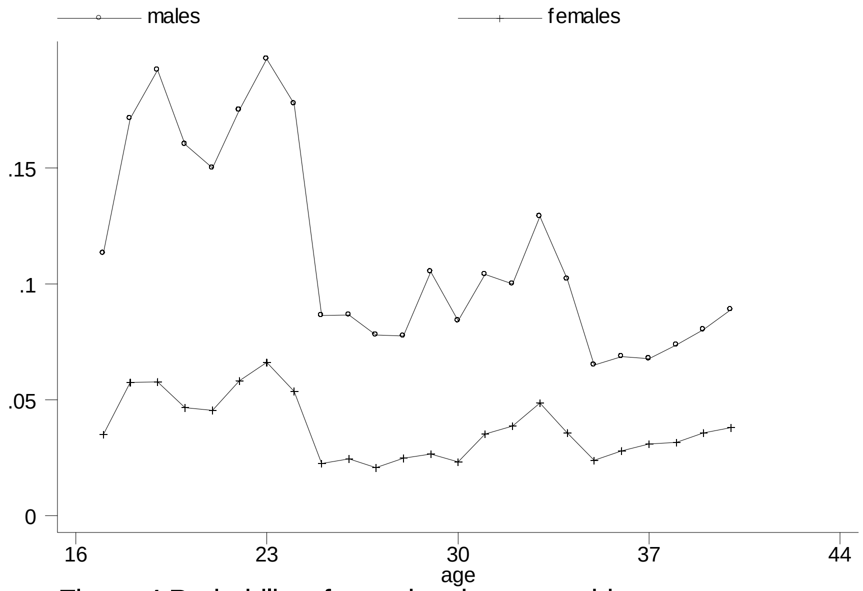


Figure 4:Probability of experiencing an accident per age

Figure 5: Observed work-disability profiles over time

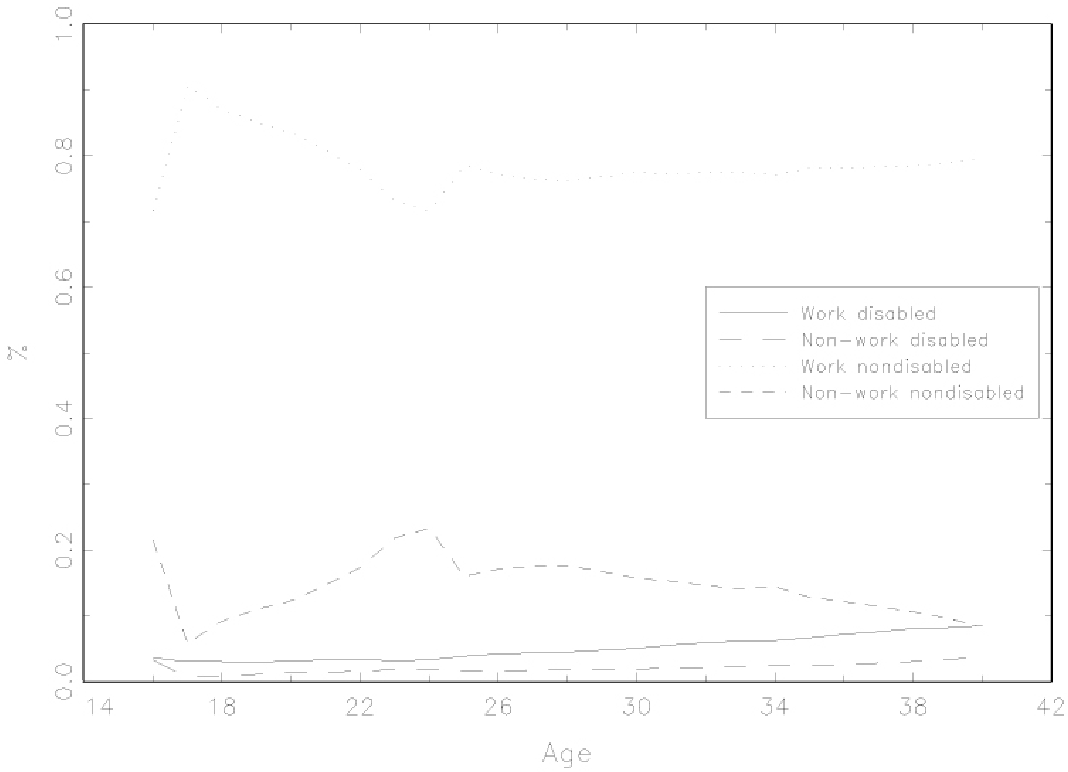
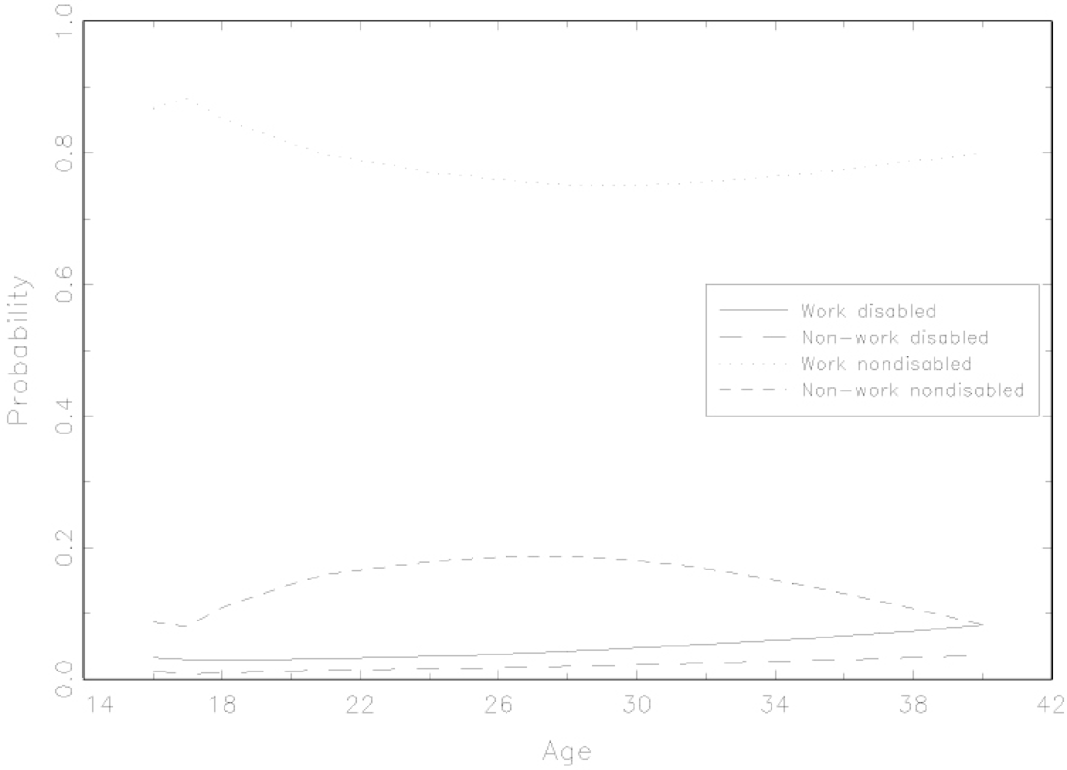


Figure 6: Predicted work-disability profiles over time



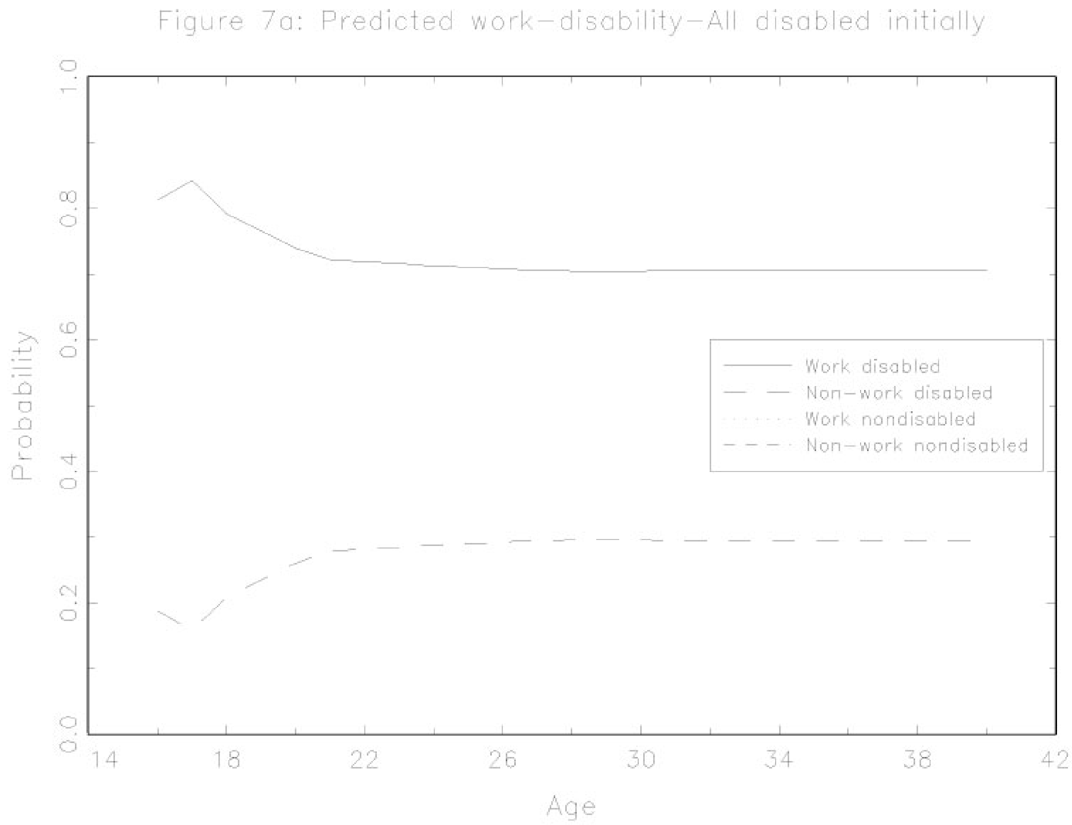


Figure 7a: Predicted work-disability-All disabled initially
Probability
Age
Work disabled
Non-work disabled
Work nondisabled
Non-work nondisabled
0.0
0.2
0.4
0.6
0.8
1.0
14
18
22
26
30
34
38
42

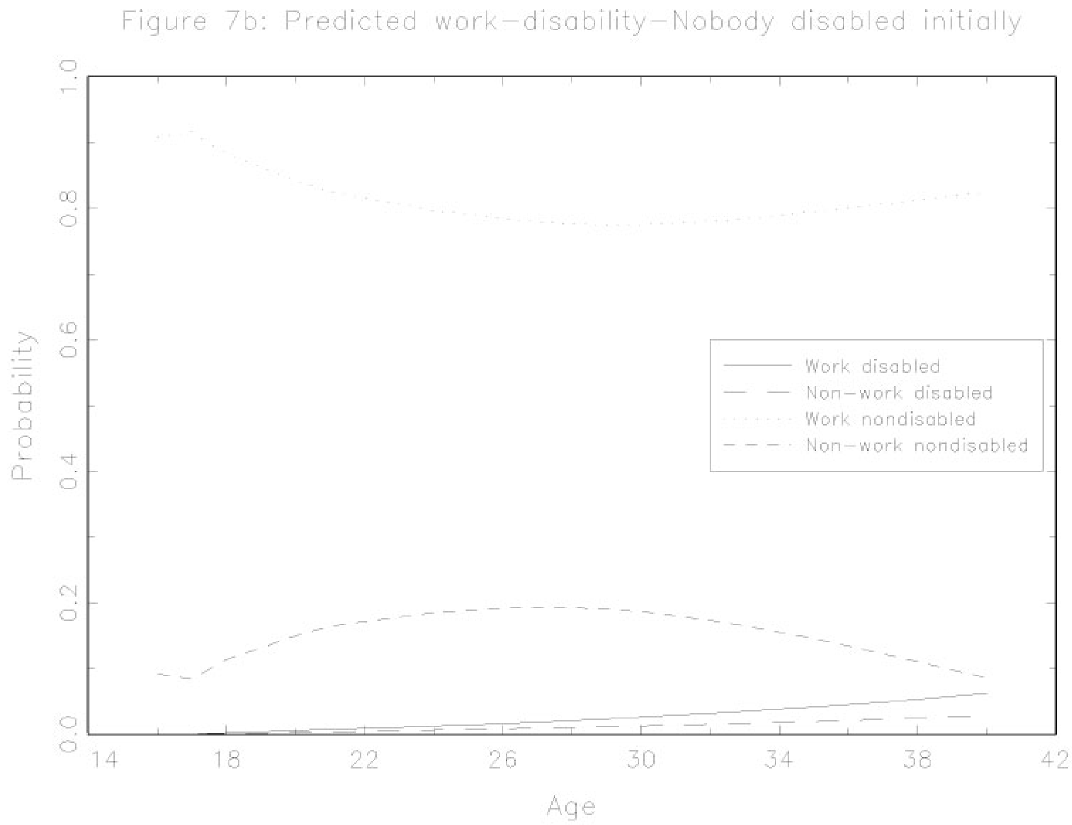


Figure 7b: Predicted work-disability-Nobody disabled initially
Probability
Age
Work disabled
Non-work disabled
Work nondisabled
Non-work nondisabled
0.0
0.2
0.4
0.6
0.8
1.0
14
18
22
26
30
34
38
42

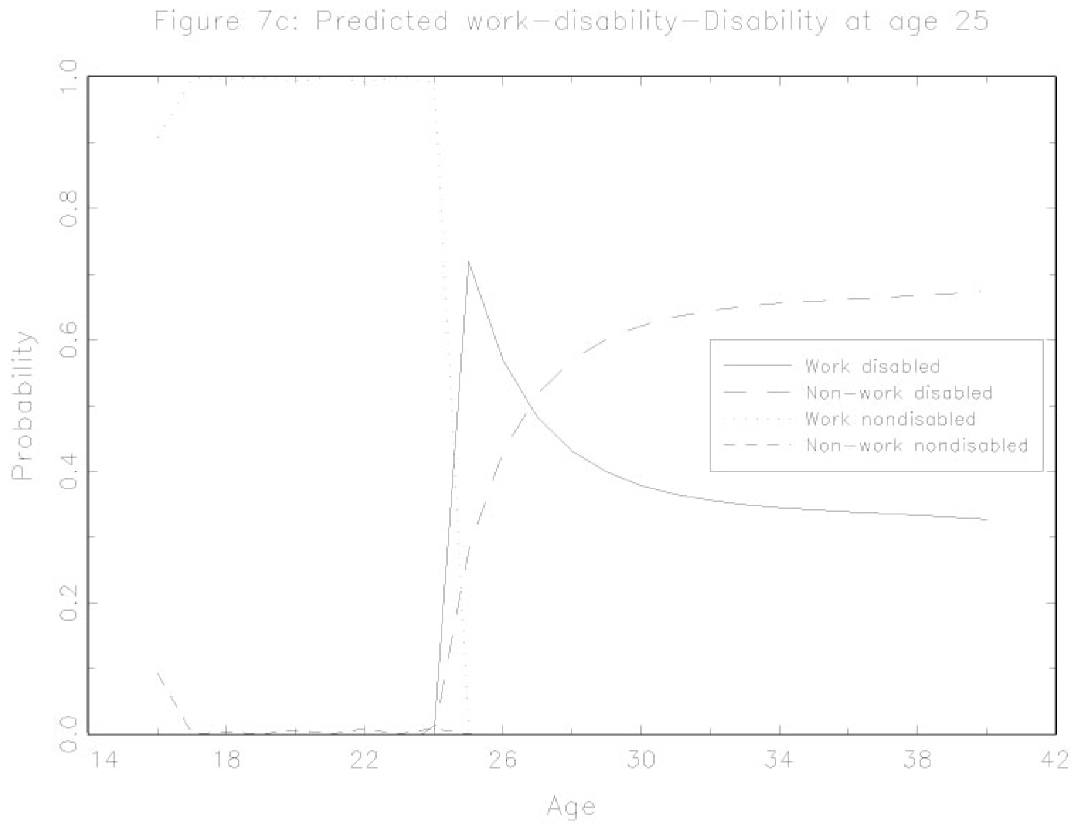


Figure 7c: Predicted work-disability-Disability at age 25
Probability
Age
Work disabled
Non-work disabled
Work nondisabled
Non-work nondisabled
0.0
0.2
0.4
0.6
0.8
1.0
14
18
22
26
30
34
38
42

Figure 8a: Predicted work-disability profiles—No accidents

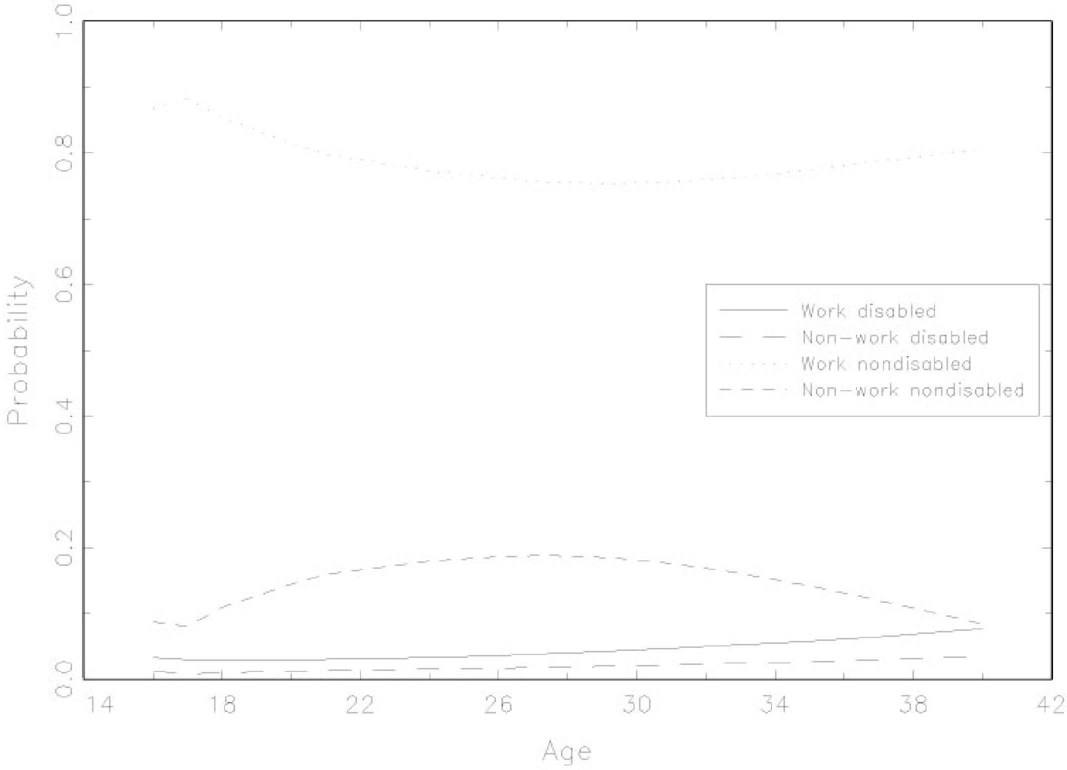
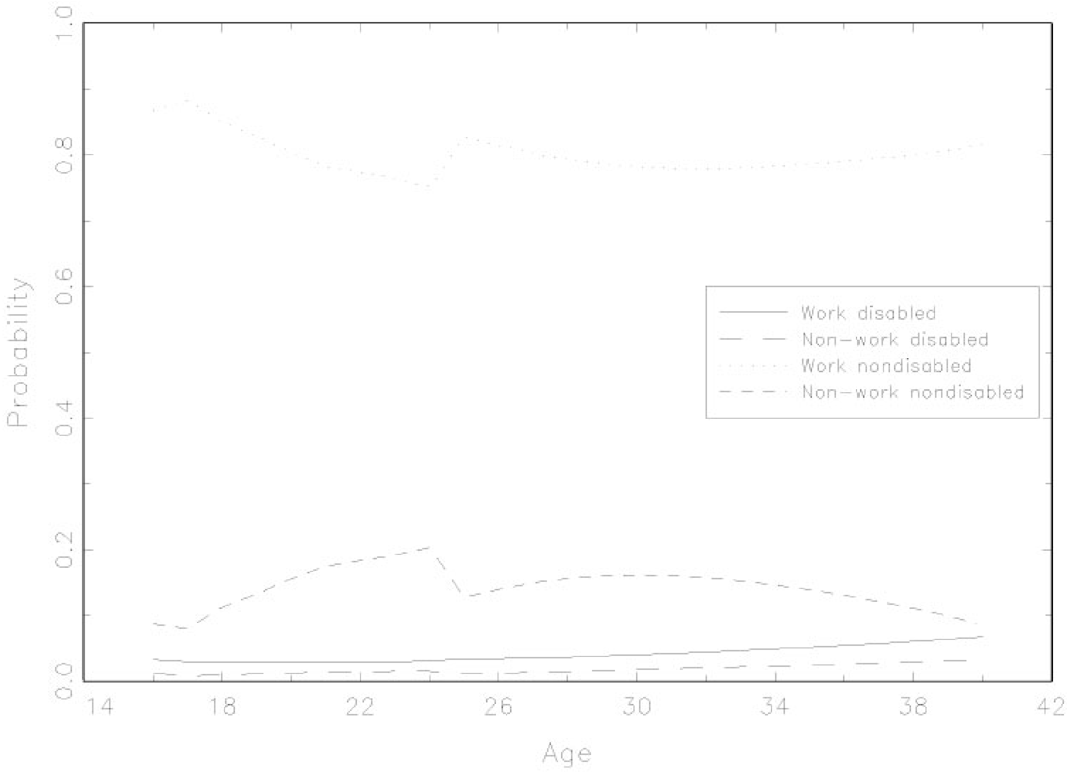


Figure 8b: Predicted work-disability profiles—Accident age 25



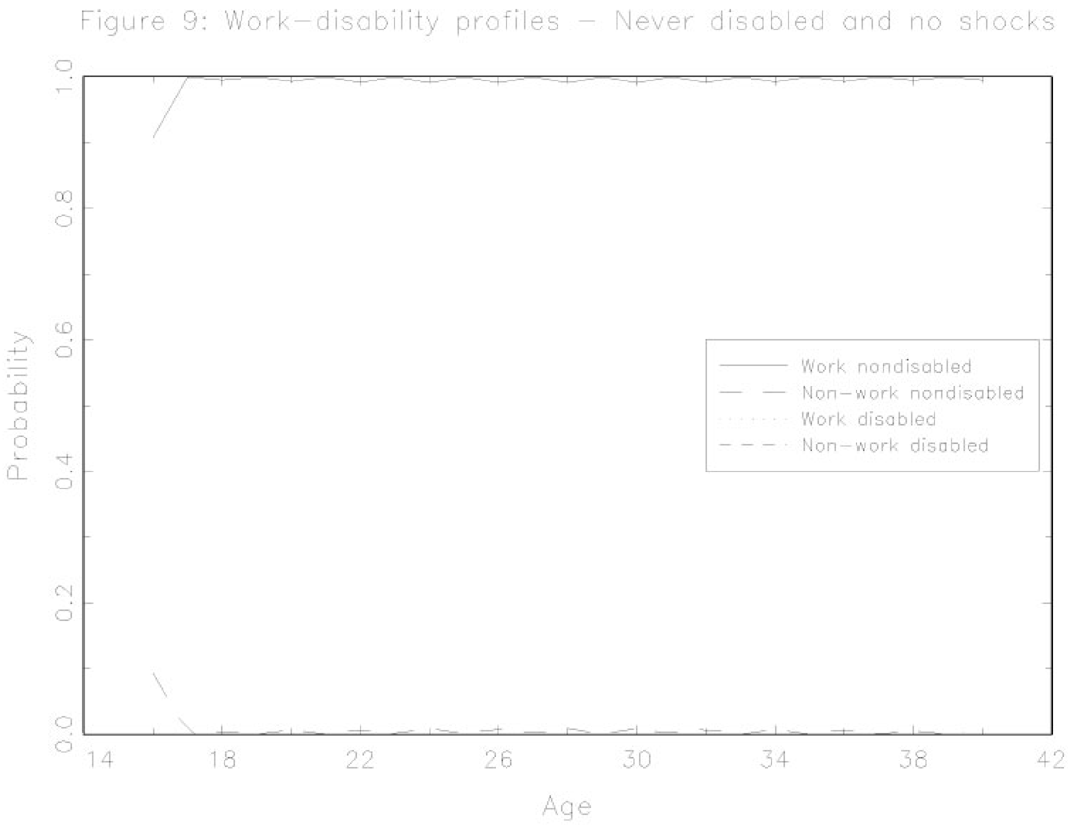


Figure 9: Work-disability profiles – Never disabled and no shocks
Probability
0.0
0.2
0.4
0.6
0.8
1.0
14
18
22
26
30
34
38
42
Age
Work nondisabled
Non-work nondisabled
Work disabled
Non-work disabled

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	