Lindeboom, Maarten; Llena-Nozal, Ana; van der Klaauw, Bas

Working Paper

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

IZA Discussion Papers, No. 2096

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Lindeboom, Maarten; Llena-Nozal, Ana; van der Klaauw, Bas (2006) : Disability and work: the role of health shocks and childhood circumstances, IZA Discussion Papers, No. 2096, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/33382

DISCUSSION PAPER SERIES

IZA DP No. 2096

Disability and Work: The Role of Health Shocks and Childhood Circumstances

Maarten Lindeboom
Ana Llena-Nozal
Bas van der Klaauw

April 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Disability and Work: The Role of Health Shocks and Childhood Circumstances

Maarten Lindeboom
Free University Amsterdam, Tinbergen Institute, HEB, Netspar and IZA Bonn

Ana Llena-Nozal
Free University Amsterdam and Tinbergen Institute

Bas van der Klaauw
Free University Amsterdam, Tinbergen Institute, Scholar, CEPR and IZA Bonn

Discussion Paper No. 2096
April 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 2096
April 2006

ABSTRACT

Disability and Work: The Role of Health Shocks and Childhood Circumstances*

This paper focuses on the relation between the onset of disability and employment outcomes. We develop an event history model that includes unscheduled hospitalizations as a measure for unanticipated health shocks and estimate the model on data from the British National Child Development Study (NCDS). We show that such health shocks increase the likelihood of an onset of a disability by around 138%. However, health shocks are relatively rare events and therefore the larger part of observed disability rates result from gradual deteriorations in health. We find no direct effect of health shocks on employment outcomes. Using the health shock as an instrumental variable shows that the onset of a disability at age 25 causally reduces the employment rate at age 40 with around 21 percentage points. Our results show that early childhood conditions are important in explaining adult health and socioeconomic outcomes. Those who have experienced bad conditions during early childhood have higher rates of health deterioration during adulthood, are more likely to become non-employed and suffer from longer spells of non-employment during the course of life.

JEL Classification: I12, J28

Keywords: disability, work, health shocks, event-history model, early childhood

Corresponding author:

Bas van der Klaauw
Department of Economics
Free University Amsterdam
De Boelelaan 1105
NL-1081 HV Amsterdam
The Netherlands
Email: klaauw@tinbergen.nl

* We would like to thank Ana Rute Cardoso, Regina Riphahn, Owen O'Donell and participants of the IZA Workshop "The Older Worker", Lisbon, 2005, iHEA world conference, Barcelona, 2005, the Ecuity III workshop in Bonn 2005 and seminar participants at Antwerp University and the Paris School of Economics for useful comments.

1. Introduction

There exists a strong positive association between health and socioeconomic status in 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 observed in later life. In this paper we focus on the role of health shocks, how important they are for disability and work outcomes and whether the relationship between shocks, disability and work varies with socio-economic background during childhood.

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 income from work and job security may affect health positively, while stress and adverse working conditions can increase the rate of health deterioration. Indirectly, because the employment status may also influence the likelihood of experiencing an adverse health shock. Whether the decline in health is gradual or falls abruptly due to shocks, it may lead to long standing disabilities that restrict individuals in doing their daily and/or work activities. This in turn may affect labor supply decisions and later work outcomes.

Smith (1999) describes the ongoing debate about the direction of the causal relations between health and socioeconomic status. In general, it is difficult to disentangle the underlying causal mechanisms, mainly because unobservables relate to both health and work outcomes. Identification of the causal relations between health and labor market outcomes requires independent variation in either health or work status to assess the effect of one on the other. Lindahl (2005), for example, uses lottery prize winning to study the effect of income on health.

In this paper we use unscheduled hospitalization as a measure for (adverse) health shocks. These health shocks are important for two reasons. First, there is a direct interest in the effects of health shocks on disability and work outcomes, how important are these shocks in explaining disability rates. Second, unscheduled hospitalizations provide unanticipated variation in health status, which can be used to identify the causal effect of the onset of disability on work status. In our context no anticipation means that the exact timing of an unscheduled hospital visit is not known in advance. This does not rule out that individuals may be aware that at some moments the risk of experiencing such a health shock is higher than in other periods, or that this risk, for instance, depends on the current employment status. In particular, a substantial share of the adverse health shocks is related to work. Also we do not require health shocks to be

exogenous, the risk of experiencing a health shock is allowed to depend on both observables and unobservables.

We construct an event history model for transitions between work and disability states and we allow the transition rates to be affected by the health shocks. The transition rates and the likelihood of experiencing a health shocks are related through unobservables. To estimate the model we use data from the British National Child Development Study (NCDS), which is a longitudinal study of around 17,000 individuals born in Great Britain in the week of 3-9 March 1958. These individuals are followed from birth up to the year 2000, when they were 42 years old. The data contain abundant information on the situation of the family where the individual was born in and early childhood health outcomes. At age 40 already about 12% of the respondents face a permanent disability and about 29% of the disabled are out of work. These numbers show that disabilities and labor outflow are already substantial at relatively young ages.

Our results show that health shocks are important for explaining disability rates, experiencing an unscheduled hospitalization increases the probability of the onset of a disability by 138%. However, because unscheduled hospitalizations are rare events, the larger part of the onsets of disabilities come from gradual deterioration of health. Our estimation results show that health shocks affect the labor market status only indirectly through the onset of disabilities. We therefore argue that an unscheduled hospitalization can be used as an instrumental variable for the onset of disabilities and that the causal effect of the onset of a disability at age 25 on the employment rate at age 40 is about -0.205. Separate estimations for males and females show that the onset of a disability has more than twice as large an effect on the employment rates of males than of females.

As a key element to the association between health and socioeconomic status during the life cycle, early childhood conditions 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 et al, 2005). Our estimation results show that people who have experienced adverse conditions during early childhood more often get adverse health shocks, start their working career in worse health and employment states and have higher probabilities of becoming disabled and non-working during their prime ages. We use our model to shed more light on the mechanisms underlying this finding. Our results show that prenatal and postnatal conditions affect later life outcomes directly primarily because they affect the rate at which health deteriorates during adult ages and the transition rates into and out of employment. In addition prenatal conditions have a strong impact on the probability to start the working career with a permanent disability.

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. Finally, Section 5 concludes.

2. Theoretical background and specification of empirical model

2.1 Theoretical Background

Health production models (e.g. Grossman, 1972, or other model based on this seminal paper) assume 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 are rational agents and according to the model they include expectations about their health trajectories when making decisions regarding health behavior and labor supply. With new information, people update their expectations and change their behavior accordingly. This underlines the difficulties in identifying the causal relations between health and socioeconomic outcomes such as labor market status. If health trajectories are predictable, individuals anticipate to that and change their behavior accordingly. So an observed change in labor market status that precedes a health transition can be the result of anticipated behavior, rather than labor market status causally affecting health.

Empirical analyses are often plagued by the presence of unobservables related to both health and socio-economic status (see for a survey of empirical studies Currie and Madrian, 1999). Only a relatively small number of studies have used panel data to control for unobservables, but even then exogenous variation is required to assess causal effects. A few have used natural experiments. Lindahl (2005), for instance, finds using lottery prize winners that the effect of income on health is significant, but rather small.

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 contains new information to the individual and thereby provides some exogenous variation in health that is unrelated to work status. Smith (2003) uses the onset of chronic conditions as a measure for health shocks and examines their effect on 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 et al (2003) conclude for a sample of older (69 plus)

individuals that these types of health shocks affect wealth only marginally.[1] Møller-Danø (2005) uses road accidents as a measure for health shocks and finds long lasting income and employment effects.

We will use unscheduled hospital visits to investigate the effect of adverse health shocks on labor market outcomes and the onset of a disability. The British NCDS 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. 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.

Since health shocks occur at different moments in life, our model should be dynamic. A dynamic model also has the advantage that we can substantially relax the requirements for unscheduled hospitalizations to be valid health shocks. Within our dynamic model we allow these health shocks to be endogenous, i.e. we explicitly model the occurrence of a health shock and allow unobservables to affect jointly the probability of experiencing a health shock, the onset of disabilities and labor market outcomes.[2] The advantage of a dynamic model is that if health shocks are unanticipated in the sense that people can not fully predict the exact *timing* of the occurrence of the shock, the effect of the 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 in the context of event history models). If one is willing to assume that the timing of the onset of a disability is also unpredictable (Smith, 1998,1999, 2003) the effect of the onset of a disability on work in our model can also be interpreted as causal. We will be more specific about our dynamic model and identification issues in Subsection 2.2.

Poor childhood health and socio-economic background are often considered to be important contributors to the association between health and socioeconomic outcomes at later ages. (e.g. Case et al, 2002, Currie and Hyson, 1999, Currie and Stabile, 2002, Dobblhammer, 2003). For instance, Currie and Hyson (1999) investigate the consequences of low birth weight. They find that children

[1] 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.

[2] Smith (1998, 1999) suggests including all possible risk factors of experiencing adverse health shocks as an alternative to a model with unobserved individual specific effects.

whose parents have a low socioeconomic status more often have a low birth weight and are less likely to recover from the negative consequences. Fuchs (2004) finds that individuals who have experienced adverse early childhood circumstances often have lower educational attainments and that the level of education is correlated with health later in adulthood. Case et al (2005) find that controlling for parental income, education and social class, children exposed to poorer uterine environments and poorer childhood health have significant lower educational attainment and poorer health and lower socioeconomic status as adults.

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 Dobblhammer, 2003 and Case et al, 2005). According to the *fetal origins hypothesis* adverse conditions during pregnancy increase disease risks in later life (Barker, 1995). So according to this hypothesis there is a direct and long lasting effect of adverse conditions during pregnancy. To the extend that a lower socioeconomic background during childhood is correlated with adverse conditions in *utero*, we may also observe that people from lower socioeconomic background have higher risks of experiencing adverse health shocks during the course of life and that their health deteriorates faster at adult ages. Alternatively, *life course models* argue that bad conditions in childhood (primarily postnatal) persists until late adulthood and may in addition influence later childhood health and schooling outcomes, health status and labor market opportunities at the start of the working career and subsequently later life health and socioeconomic status. *Pathway models* (Marmot et al, 2001) argue that the negative effect of adverse early childhood is primarily via teenage health and health and socioeconomic outcomes during early adulthood. So, these models hypothesize that socioeconomic status and health in early adulthood are the most important determinants of later life health and that there is no or hardly any effect of early childhood conditions after that we have conditioned on these. Case et al (2005) find, using the same data as we use in our analyses, that all three models play a potentially important role. Uterine environment (measured by low birth weight and whether the mother smoked during the pregnancy) and childhood health have a significant and lasting impact on health and socio-economic status in middle age. Childhood factors (measured by socioeconomic position of the parents) are important determinants of social status during early adulthood and also influence later life health directly..

In our analyses we will capture early childhood conditions with a set of variables that cover elements of prenatal health and childhood health and socio-economic background. We will be more specific about the precise definition of these variables later. In our model we will allow early

childhood conditions (prenatal and postnatal) to affect health and labor market outcomes in three possible ways. First, we allow for a direct effect of early childhood conditions on disability and labor market outcomes in early adulthood (i.e. at the moment of leaving school). The latter outcomes are allowed to affect disability and labor market outcomes at later (middle) ages. If the hypothesis of Marmot et al (2001) would be true, then this would be the only relevant effect of early childhood conditions. Second, early childhood conditions are allowed to directly affect the rate of health depreciation and the incidences and length of work(less) spells during middle ages. Persistent effects of prenatal variables, after we have controlled for health and socio-economic status during early adulthood, hints at the relevance of the fetal origins hypothesis. Persistent effects of postnatal childhood variables hints at the relevance of life course models. This is in agreement with what Case et al (2005) do in the context of their linear model. It has to be noted however, that it will always be difficult to make a strict distinction between the two hypotheses. The set of variables that we use for prenatal conditions may be incomplete and part of this may be captured by the set of childhood variables. Finally, we allow the probability of experiencing an adverse health shock during the course of life to depend on early childhood conditions. This implies that adverse childhood conditions may be a trigger for later health shocks, which in turn may influence disability and labor market outcomes during adulthood. In Section 4 we turn to the relative importance of these three effects for disability and work outcomes in middle age and a discussion of the relevance of our findings for the different hypotheses implied by the literature.

2.2 Empirical specification

In this section we describe our empirical model, but first we briefly sketch the structure and contents of our data. We observe individuals from birth up to the age of 42 and have constructed individual labor market histories since the moment the individual leaves full-time education. The labor market histories contain yearly information on employment status (employed or non-employed) and disability status. We only focus on permanent disabilities and thus ignore short-term limitations. Finally, for each year we observe whether there was a major health event that lead to a hospitalization, and whether this was scheduled or not. In our model we use unscheduled hospitalizations, labeled below as health shocks. In the next section we discuss the data in more detail and return to the definition of the labor market states, disabilities and health shocks.

We use a discrete-time event history model to analyze transitions between different disability and work states. The model is a semi-Markov model that contains 4 states. Let $S_i(t)$ denote the individual's labor market status at the beginning of year t, this can either be working (1) or non-working (0). In each year the individual 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 year *t*, which can either be disabled (1) or non-disabled (0). Because we only focus on permanent disabilities, being disabled is an absorbing state, a once disabled individual cannot recover. The transition probabilities for moving between different states are affected by health shocks that might occur to the individual. The variable $A(t)$ takes the value 1 if a health shock occurred between year *t* and $t+1$ and 0 if no health shock occurred in this year. The probability of experiencing a health shock is allowed to depend on the individual's current labor market status as health shocks can be work related. The probability of a health shock between *t* and $t+1$ equals:

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

The transition probabilities between the different disability and work states 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, for the accident probability *q*:

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

where x_t is a vector of the individual's socioeconomic characteristics (including an intercept) at time *t* and v_a is an unobserved individual component that does not vary over time. The parameter δ describes effect of being employed on the risk of having an adverse health shocks. The transition probabilities are specified as:

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

if $(i,j)\neq(k,m)$ and

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

The parameters $\eta_{(i,j),(k,m)}$ describe the effects of health shocks on the different transition probabilities. The impact of experiencing an adverse health shock can thus be different for individuals in different work and disability states.

The transition probabilities and the probability of having a health shock 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 components 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_n denotes the probability that the elements in w_n equals 1. The unobserved heterogeneity term follows

$$v_a = w'\alpha_a$$

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 elements as the vector w and are estimated along with the other model parameters.

Since, the model is fully parameterized, we can use maximum likelihood to estimate all parameters. Therefore, we use for an individual who we can follow for T years, the sequences of labor market, health states and health shocks given by $s_l(1)$, $s_l(2),\ldots, s_l(T)$, $s_h(1)$, $s_h(2),\ldots, s_h(T)$ and $a(1),a(2),\ldots,a(T)$, respectively. In the estimation we condition on the initial labor market status and health status of the individual (when this individual leaves full-time education) as given. In Section 4 we will estimate a multinomial logit model for these initial states and investigate the sensitivity of the initial state to early childhood conditions.

The first set of parameters of interest are those describing the effect of a health shock on disability and work outcomes, i.e. the parameters η. Whether this effect can be interpreted as causal hinges on the assumption that individuals can not anticipate the exact moment of the arrival of the health shock. This does not imply that health shocks are exogenous or that each

individual has in each time period the same probability of experiencing a health shock. The probability that health shocks occur can differ between individuals, based on both observed and unobserved characteristics. Furthermore, individuals might know that in particular periods the probability of getting a health shock is high, for example when they are employed or as they get older. We only assume that in advance individuals do not know the exact timing of a health shock. This assumption is satisfied by the definition of a health shock as an unscheduled hospital visit. See Abbring and Van den Berg (2003) for an extensive discussion on identifying the effects of unanticipated interventions in dynamic models.

To identify the causal effect from the onset of a disability on employment outcomes, we can use two alternative strategies. Like e.g. Smith (1998, 1999, 2003) we could assume that also the onset of a disability is an unanticipated health shock, which would identify the causal effect on employment outcomes along the same line of reasoning as above. However, in our application we have the health shocks as alternative variation and the empirical results show that health shocks mainly affect the onset of disabilities. Within our dynamic model, we could thus use the health shocks as instrumental variable for the onset of disabilities when measuring the causal effect of the onset of disabilities on employment outcomes. We return to this issue is Section 4 when we discuss the estimation results and the consequences.

Finally, we are interested in disentangling the association between early childhood conditions and health and socioeconomic outcomes during adulthood. Our data contain a number of indicators for early childhood conditions, such as the socioeconomic background of the parent and birth weight. These variables are included in the vector x_t. Without making strong exogeneity assumptions, we cannot identify the causal effect of each indicator. But we can identify whether the indicators for poor early childhood conditions are jointly important in explaining the occurrence of health shocks, the transitions between states and/or the initial state after leaving full-time education.

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 1999/2000 (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 were 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 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 health shocks 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.[3] After selecting only those with complete labor and health histories, our final sample consists of 12,448 individuals. Case et al (2005) 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

[3] 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)

(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, the *p*-values for joint significance of these two variables the 1991 and the 1999/2000-wave are both very close to 0 (results not included). In particular, employed individuals and disabled are more likely to participate in later waves. Of relevance is therefore whether the parameter estimates of our statistical model are sensitive with respect to this attrition. Therefore, as a sensitivity analysis we have also estimated the model of Section 2, where we include in the set of explanatory variables a dummy variable indicating whether or not an individuals dropped out of the panel before the last wave. These dummy variables are significant in explaining transition rates, but do not change the estimates of our parameters of interest (results are available on request).

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 males 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 increases with age. The unemployment rate is relatively constant except for the ages 22 - 24, unemployment is somewhat higher for these ages. This might be related to a business cycle effect, i.e. the recession in the late 1970s/beginning 1980s. 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, the fraction of females out of the labor force starts to drop and 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 disabilities such as epilepsy, blindness, deafness, multiple sclerosis, mental retardation, a congenital condition, or a traumatic amputation or internal injury. In the Appendix 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-9 is extensively used in epidemiological and health management studies to classify diseases and health problems (World Health Organization, 2004). Case et al (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 rates for employed workers 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 20 over 4% of the individuals in the sample has some disability. This increases up to about 12% at age 40. 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 the onset of a disability becomes larger as people get older.

In this paper we define a health shock as an 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 health shock and the type of health shock.[4] Men are much more likely to experience health shocks than women. In our sample, around 77% of the men had at least one health shock during the observation period, while this was only about 42% for women. Multiple health shocks for a single individual are frequently observed. Not only the incidence of health shocks differs between men and women, but also the types of health shocks differ. Table 1 lists the annual incidence rates for different types of health shocks. For each type of health shock men are much more likely to experience this than women. The most substantial difference in incidence rates occurs for work and sports-related health shocks. Because, a large share of the health shocks are work related, it is particularly important in our empirical model to take account of the labor market status of the individual when we specify our model for health shocks. Figure 4 plots annual incidence rates by age and shows that for both men and women the

[4] The questionnaire restricts the number of health shocks 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.

probability of getting a health shock 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, non-work and disabled, work and non-disabled and non-work and non-disabled. 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 the non-work non-disabled state decreases while the fractions of individuals increase in both disabled states (either with or without work). 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 et al (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 et al (2005) by constructing dummy variables that indicate if the information on a variable is missing.

Low birth weight is a dummy variable for infants with a birth weight below 2500 grams. There is evidence from the epidemiological literature 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 poor 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. 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. The math test is designed for the NCDS and assesses arithmetic ability. The score ranges from 0 to 10. 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 Bristol Social Adjustment Guide, is designed to assess the child's social behavior in school and at home. The test is completed by the teacher who knows the child best. Higher scores indicate higher maladjustment. The data also include information on the Southgate Reading Test. However, including this test score in our analyses did not improve our empirical results after inclusion of the math score and Bristol Social Adjustment Guide. Therefore, we decided not to use this reading test score in the analyses.

The education level is derived 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. Finally, we will use the region at birth to control for geographical differences and/or differences in labor market conditions.

4. Empirical results

We start with a brief discussion of the results for the mixing distribution (unobserved heterogeneity). The parameters of the mixing distribution are reported in the lower panels of Tables 4a and 4b. The preferred specification is a factor-loading 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). Therefore, for each transition probability there are four mass point. Most probability mass is located at a mass point 3, describing individuals with a low probability of experiencing a

health shock (see Table 4a). Individuals who are most likely to get a health shock (mass point 2 in Table 4a) are also more likely to switch states (see the values of the mass point 2, compared to the other mass points in Table 4b). The other two mass points describe individuals who have an average probability of experiencing a health shock, but are either not very likely to switch labor market and disability status (mass point 4) or are much more likely than other individuals to change states (mass point 1).

Concerning the main parameters of interest, Table 4a shows the parameter estimates from the logit specification for the probability of getting a health shock. Employed individuals have about a 45% (=exp(0.371)-1) higher probability of getting a health shock. Recall from Table 1 that indeed a substantial share of the health shocks are work related. Males are about three times more likely to get a health shock than females. This is also what we directly observe in the data when we compare males and females. Consequently, gender differences in employment rates and observed individual characteristics cannot explain the differences in health shock incidences between men and women. The probability of getting a health shock is U-shaped in age, it is decreasing until age 38 and increasing afterwards. Health at birth and cognitive ability during childhood years are important. In particular, individuals whose mother smoked during pregnancy are more likely to suffer from adverse health shocks and the probability of having a health shock increases with the mother's age at birth. The parental socioeconomic status also has a significant effect on the rate at which health shocks occur. Early childhood conditions are thus important in explaining adverse health shocks during adulthood. Height at age 23 is important, taller people have more health shocks. Individuals with a high math score at age 7 and who were less socially adjusted (high values of the Bristol Social Adjustment Guide score) also have higher probabilities of getting a health shock. It is difficult to connect a strong causal interpretation to these findings since, 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 health shocks.

Table 4b shows the parameter estimates of a multinomial logit model for the transitions between the different labor market and disability states. Of central importance are the effects of health shocks on transitions. These effects are summarized in Figure 6. The thick arrows are associated with large coefficients. The figure reveals that, as expected, health shocks primarily have an effect on the transition rates from non-disabled to disabled. It is difficult to interpret the coefficients separately from each other. To illustrate the impact on a health shock we therefore consider a representative individual.[5] The probability that this individual is non-disabled at his

[5] In fact we simulate the model for all individuals and compute average (transition) probabilities over all individuals.

24th birthday is 0.952. Without experiencing a health shock at age 24 the probability of becoming disabled before his 25th birthday is 0.0030. However, if the individual actually experienced a health shock, this probability becomes 0.0070. Experiencing a health shock thus increases the instantaneous onset of a disability with around 138%. It should be noted that the health shock at age 24 causes a lasting difference in disability rates. If this individual does not suffer from other health shocks anymore, then with a health shock at age 24, the disability rate at age 40 is 0.1173. The disability rate at age 40 equals 0.1137 if the individual never gets a health shock.

The direct effects of a health shock on employment rates are negligible. To illustrate this, we consider the representative individual who is at his 24th birthday non-disabled. Again, we compare the situation where this individual does not experience any health shock, with the case that this individual gets a health shock at age 24. But we impose that the health shock does not cause the onset of a disability, i.e. at his 25th birthday the individual is still non-disabled. In this case employment rates are unaffected by the health shocks. If for example, the representative individual is still non-disabled at age 40, then the probability of being employed is 0.9013 regardless of having experienced the health shock.

Indeed it is difficult to come up with convincing stories for effects of health shocks on employment, other then via an effect on disability/health. This implies a process where health shocks may trigger the onset of disabilities and subsequently disabilities may affect employment status. Within this framework, a health shock can be seen as an instrumental variable for the causal effect of a disability on employment status. To compute the instrumental variable estimator for the causal effect of the onset of a disability on employment status, we again consider the representative individual. Without any health shock before age 40, his employment rate at age 40 is 0.8778. If this individual experiences a health shock at age 24, then the employment rate at age 40 equals 0.8770. The Wald estimator for causal effect of becoming disabled at age 25 on the employment status at age 40 is

$$\frac{\Pr(S_1(40)=1 \mid A(24)=1) - \Pr(S_1(40)=1 \mid A(24)=0)}{\Pr(S_h(25)=1 \mid A(24)=1) - \Pr(S_h(25)=1 \mid A(24)=0)} = -0.205$$

This implies that the onset of a disability at age 25 causes a reduction in the probability of being employed at age 40 by 0.205.

Usually in the economic literature that focuses on the relation between disabilities and socioeconomic outcomes, the onset of a disability is assumed to be an unanticipated event (e.g. Adams et al, 2003, and Smith, 1998, 1999, 2003). When we use our model to simulate the

employment effect of an onset of a disability at age 25 compared to not getting a disability.[6] The difference in employment rates ate age 40 is 0.229, which is close to the Wald estimator above.

We previously assessed that health shocks are important for the onset of disabilities and that the effect is lasting. However, until age 40 men experience on average about 2.4 unscheduled hospitalizations and women only 0.8. So the occurrences of shocks are relatively rare events. Calculations with the model show that they can account for only about 6.6% of all disabilities at age 40. Hence, the larger part of long standing disabilities arises from a gradual deterioration in health.

Being female increases the transition rate from the employment state towards non-employment 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-working 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 (although it should be noted that women experience less health shocks). Non-working women have lower probabilities of becoming disabled than non-working men.

We have estimated separate models for males and females (parameter estimates available on request). Again we find that health shocks have significant impacts on disability rates. The occurrence of a health shock increases the probability of an onset of a disability with 95% for women and 154% for men. Women are thus not only less likely to experience health shocks, but also if they get a health shock, the impact is less severe. At age 40 men and women have equal disability rates. Therefore we can conclude that the gradual deterioration in health is more important for females. For both men and women health shocks do not have a direct effect on employment rates. The Wald estimator of the causal effect of a disability on employment is for males -0.234 and for women -0.093. The onset of a disability thus has a larger effect on the employment rates of men than of women. Differences in occupational choices between men and women may be important for this. Also in our analyses we do not make a distinction between part-time and full-time work. Females are more often employed in part-time jobs and it may be easier to continue working in these jobs after the onset of a disability.

From Table 4b we see that after age 20 the probability to get a disability increases. Furthermore, parental socioeconomic status, whether the mother smoked during pregnancy, mother's age at birth and the indicator for low birth weight have significant effects on almost all transition rates. A general picture that emerges from these coefficients is that adverse early childhood circumstances increase the probability of becoming disabled, the incidence of entering

[6] So we use $Pr(S_l(40)=1|S_h(25)=1)$-$Pr(S_l(40)=1|S_h(25)=0)$.

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.

We made some calculations with the model to make the results of early childhood conditions more insightful. In particular, we consider two representative individuals with similar characteristics except for the parental socioeconomic status at birth. The first individual comes from parents with a high socioeconomic status and the second individual from parents with a low socioeconomic status. Average disability rates are 5.7% at age 25 for the individual from a low socioeconomic status and 4.8% for the individual from a high socioeconomic status. At age 40 the disability rates are 13.5% and 10.4%, for individuals from a low and a high socioeconomic background, respectively. In Figure 7 we show the employment rates conditional on disability status for both individuals. The figure shows that for a given disability status employment rates are higher for individuals from a high socioeconomic status.

We have estimated our model separately for individuals from a high and a low socioeconomic background. Like Currie and Stabile (2003), we do not find differences in the long-term effects of the onset of chronic conditions for these groups. In particular, the Wald estimator for the causal effect of the onset of a disability on employment is -0.180 for individuals from a high socioeconomic background and -0.183 for individuals from a low socioeconomic background.

We have seen in Table 4a that early childhood conditions affect the rate at which health shocks occur, this holds for some of the prenatal variables (whether the mother smoked during the pregnancy) and for most of the postnatal variables (parental socioeconomic status and the test scores at age 7). The significance of the prenatal variables may indicate the relevance of the fetal origins hypothesis and the significance of the postnatal variables may point at the relevance of the life course model. However, as noted before in Section 2, it is difficult to empirically distinguish between the two hypotheses. The set of prenatal variables may not capture all of the prenatal conditions and part of this may be picked up by the postnatal variables.

Health shocks only explain a very small fraction of all disabilities, which limits the relevance of these indirect effects. However, we also find strong direct effects of the childhood variables (prenatal and postnatal) on the transition rates between the disability and employment states (Table 4b). However, the transition model did not include education as an explanatory variable. Currie and Hyson (1999) find that the effects of early childhood conditions (measured by low birth weight) are largest on educational attainments. In light of these result it is interesting to see if childhood conditions persist in adult age, after we condition on educational attainment.

Estimation results with the model including the education level as a regressor show that even though the education level has a significant impact, early childhood conditions remain important factors for disability and work transitions during adult ages (estimation results available on request). Calculations with this model show that early childhood conditions have a sizable effect on employment rates and disability rates. The differences in disability rate is still 0.7% points at age 25 (it was, see above, 0.9 % points) at age 40 this difference is now 2.1% (was 3.1%, see above) points. Also the differences in employment probabilities (conditional on disability status) between individuals from high and low socioeconomic statuses become somewhat smaller after controlling for the level of education. However, these differences remain substantial. This shows that that early childhood conditions remain important in explaining the rate of health depreciation during adulthood after controlling for early adulthood health and socioeconomic outcomes and educational attainments. The pathways models (Martmot et al 2001) states that early childhood conditions affect adult health and labor market outcomes mainly via early adulthood socioeconomic outcomes. Our result can not support this hypothesis.

To further investigate the mechanisms underlying the impact of the socioeconomic status at birth on later disability and employment outcomes we perform some simulations with our model. We focus on the marginal effects of socioeconomic status on disability and work outcomes in adult ages via (1) the initial state that individuals are in after leaving school, (2) the prevalence of health shocks and (3) the transition rates after that individuals have entered the labor market. For these simulations we in addition estimated a (multinomial logit) model for the initial state after leaving school. The results of this model are reported in Table B1 of the appendix. We first consider a representative individual, with high parental socioeconomic status at birth. The solid line in Figure 8 shows the predicted disability rate for this individual at different ages. Next, we assume that for predicting the initial state the individual comes from a low socioeconomic status at birth, while in the models for the occurrence of health shocks and transitions we maintain a high socioeconomic status. So this exercise measures the pure partial effect of socioeconomic status via changes in the initial status after leaving school. We also perform simulations where we change the socioeconomic status in the health shocks model and in the transition model. As we can see from Figure 8, the simulated disability rates only diverge substantially from the model prediction in case the socioeconomic status at birth is switched to low in the transition model. So, the figure shows that while socioeconomic status of the parents has a significant effect on the initial state and the occurrence of a health shock, the size of these effects are modest. Concerning the effect of low socioeconomic status via higher transition rates, from the figure it becomes clear that the negative relationship between socioeconomic

background and disability increases over age. Case et al (2002) already find for younger children that the health of children from lower socioeconomic backgrounds deteriorates faster. In Figure 9 we show the simulations for the employment rates. Again we see that employment rates diverge most if in the transitions rates high socioeconomic status at birth is changed to low socioeconomic status.

We repeated the simulations with the indicators for the prenatal conditions. In particular we considered a representative individual, whose mother did not smoke during the pregnancy and who had a normal birth weight. Next we repeat the calculations, but now assume that the mother smoked and that the individual had a low birth weight.[7] Again, we change the prenatal conditions separately for the initial state model, the occurrence of health shocks model and labor market and disability transitions model. The results of these simulations are reported in Figures 10 and 11. Figure 10 shows us that there are two important channels for the prenatal conditions to affect disability rates, via the initial state and via the transition rates. The large effect via the initial state is due to the strong and large effect of low birth weight on already being disabled when leaving school (See Table B1). Since a disability is a permanent condition, the effect of prenatal variables via the initial state on the disability rate remains roughly constant over age. The effect of the prenatal variables on disability rates via the transition rates is increasing with age. Figure 11 depicts the effect of the prenatal conditions on employment rates. The prenatal conditions do not have a strong effect on the employment status immediately after leaving school (Table B1), but they do have a strong effect on the transition rates. This is reflected in Figure 11; the adverse prenatal conditions affect employment rates mainly via its effect on the transition rates after individuals have left the labor market.

5. Conclusions

The focus of this paper was on the relation between disability and work in middle ages and the role of health shocks and socioeconomic background. We have developed an event-history model that describes transitions between disability and work states and we allowed these transitions to be affected by health shocks and other background characteristics. We define a health shock as an unanticipated event after which an individual is admitted to a hospital or attending a hospital outpatient or casualty department. The unanticipated nature of such a health shock is important as this allows us to identify the causal effect of the shock on disability and subsequently the causal

effect of disability on work. Furthermore, with our model we can assess the importance of these sudden health shocks for disability and work outcomes at later ages.

The empirical results show that the occurrence of an unscheduled hospitalization increases the instantaneous likelihood of the onset of a disability with about 138%. Health shocks do not have a direct effect on employment rates. A health shock is a relatively rare event so that the larger part of the age related disability increases come from a gradual deterioration of health, despite the large instantaneous effect of the health shock on the disability rate. More specifically, only about 6.6% of the disabilities at age 40 can be explained from health shocks. It has to be noted that we use a strict definition of a health shock. We look at acute changes in health due to events. Diagnosed conditions during scheduled meetings with a physician are not counted as health shocks, but labeled as gradual. We find that the onset of a disability at age 25 causally reduces the employment probability at age 40 by about 0.205. This effect is larger for males than for females (0.234 and 0.093, respectively), but we do not find any difference by socioeconomic background (0.180 for individuals from a high socioeconomic background and 0.183 for individuals from a low socioeconomic background). The latter finding is consistent with Currie and Stabile (2003).

We find lasting effects of childhood circumstances on disability and work outcomes at later ages, i.e. individuals from a low socioeconomic background have higher disability rates and lower employment rates. These long lasting effects remain, even after controlling for educational outcomes and the labor market and disability status just after leaving full-time education (labeled as the initial state). It is difficult to empirically distinguish between the different mechanisms that may relate childhood conditions and later life outcomes. After all, we do not have independent variation in childhood conditions and it is a priori not clear that we can fully separate the effects of prenatal conditions (measured with an indicator for low birth weight and whether the mother smoked during the pregnancy) from the postnatal variables (measured e.g. with socioeconomic status of the father). With our model we can, however, assess how important changes in the initial state are for later life outcomes and whether there is a large role for prenatal and postnatal variables in this. Likewise, we can repeat the exercise for health shocks and transition rates. Our results show that prenatal and postnatal conditions affect later life outcomes directly because they affect the rate of health deterioration during adult ages and the transition rates into and out of employment. In addition prenatal conditions have a strong impact on the probability to enter the labor force with a permanent disability. So, adverse prenatal conditions affect later life disability outcomes in two ways: at the start of the career and via the rate at which health deteriorates after

[7] This is in accordance with what Case et al (2005) do when they consider prenatal conditions.

individuals have entered the labor market. Already at young ages there is a strong positive association between socioeconomic status and health. Therefore, another consequence of our findings is that the association between health and socioeconomic status becomes stronger as people age. This is in line with previous findings, Case et al (2002) and Currie and Stabile (2003). The above suggests potentially important effects for the life course and the fetal origins hypothesis and less so for the pathways model. Also, our conclusions have important policy implications. Reductions in later life inequality in health and reductions in disability rates can effectively be reduced with early interventions. Our results suggest that this holds in particular for programs aimed at the prevention of low birth weight. Labor market outflow rates are substantially higher for people with a lasting and permanent disability and a substantial share of these people already have left the labor market at relatively young ages. This suggests that the effect of these interventions is not limited to health outcomes, but that there are also potentially important effects on labor market outcomes.

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 Children's Bureau*, National Child Development Studies Working Paper, Fourth Follow-up: No. Paper 11, 1983.

Bajekal M, Harries T, Breman R, Woodfield K. Review of disability estimates and definitions. *Department for Work and Pensions*, IAD Social Research Division In-house Reports: 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.

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

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* 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, Madrian BC. Health insurance and the labour market, in Ashenfelter O, Card D, eds., *Handbook of Labor Economics*. Volume 3C. Amsterdam: Elsevier North Holland,, 1999, 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. *National Bureau Economic Research, Inc.,* National Bureau Economic Research Working Papers*:* No 6943, 1999.

Doblhammer G. The late life legacy of very early life. *Max Planck Institute for Demographic Research, MPIDR,* Working Paper: No. 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.

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.

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.

Møller-Danø A. Road injuries and long-run effects on income and employment, *Health Economics* 2005; **14(9)**: 955 -970.

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.

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, Inc.,* National Bureau Economic Research Working Papers*:* No. 10063, 2003.

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.

World Health Organization. *International Classification of Diseases*, Ninth Revision. Geneva: WHO Publications, 1977.

World Health Organization. *Low birthweight*. Geneva: WHO Publications, 2004.

Table 1: Yearly incidences of different types of health shocks

	Male	Female
Overall	0.1199	0.0391
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			
	state in year t+1			
state in year t	work/ disabled	nonwork/ disabled	work/ nondisabled	nonwork/ Nondisabled
work/disabled	95.3%	4.7%		
nonwork/disabled	16.8%	83.2%		
work/nondisabled	0.3%	0.1%	96.8%	2.8%
nonwork/nondisabled	0.3%	0.7%	41.9%	57.2%

	Female			
	state in year t+1			
state in year t	work/ disabled	nonwork/ disabled	work/ nondisabled	nonwork/ Nondisabled
work/disabled	90.3%	9.7%		
nonwork/disabled	12.8%	87.2%		
work/nondisabled	0.3%	0.0%	91.7%	7.9%
nonwork/nondisabled	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 the probability of health shocks

	Parameter estimates
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)
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 (0.004)
Missing	0.046 (0.009)
Bristol Social Adjustment Guide at age 7	0.764 (0.006)
Missing	-0.082 (0.012)
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)
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)

Standard errors in parentheses

See table 4b for the number of observations and the value of the log likelihood function

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

From	Disabled				Nondisabled									
	Work		Nonwork		Work				Nonwork					
To	Disabled				Disabled		Nondisabled		Disabled				Nondisabled	
	Nonwork		Work		Work		Nonwork		Work		Nonwork		Work	
Intercept	-2.321	(0.005)	-2.111	(0.006)	-6.797	(0.004)	-7.242	(0.006)	-3.402	(0.004)	-3.159	(0.004)	-3.768	(0.005)	3.082	(0.010)
Health shocks	-0.151	(0.003)	0.154	(0.003)	0.816	(0.003)	1.444	(0.003)	0.064	(0.010)	0.739	(0.011)	0.864	(0.003)	0.190	(0.005)
Female	0.794	(0.008)	-0.447	(0,034)	0.294	(0.005)	0.864	(0.003)	0.961	(0.008)	-1.184	(0.004)	-0.689	(0.004)	-0.867	(0.005)
Age (divided by 10)	-0.180	(0.004)	0.645	(0.005)	-0.111	(0.005)	-1.429	(0.006)	2.610	(0.004)	-0.764	(0.005)	-1.035	(0.003)	-2.244	(0.006)
Age squared (divided by 100)	-0.039	(0.004)	-0.161	(0.004)	0.106	(0.004)	0.314	(0.005)	-0.598	(0.003)	0.105	(0.012)	0.246	(0.005)	0.340	(0.004)
Parental socioeconomic status at birth																
Missing	0.199	(0.004)	-0.066	(0.004)	0.130	(0.006)	0.172	(0.007)	0.282	(0.003)	0.135	(0.003)	-0.408	(0.011)	-0.153	(0.003)
High	-0.198	(0.003)	-0.102	(0.011)	-0.184	(0.007)	-0.511	(0.003)	-0.157	(0.004)	0.375	(0.003)	-0.165	(0.004)	0.210	(0.005)
Low	0.223	(0.003)	-0.126	(0.017)	0.187	(0.007)	0.215	(0.003)	0.246	(0.006)	0.456	(0.008)	-0.174	(0.003)	-0.146	(0.004)
Mother smoking at pregnancy	0.703	(0.004)	-0.017	(0.009)	0.191	(0.006)	0.395	(0.003)	0.179	(0.007)	-0.167	(0.004)	0.172	(0.003)	-0.040	(0.006)
Missing	0.158	(0.003)	-0.501	(0.004)	0.073	(0.006)	0.540	(0.003)	0.221	(0.004)	-0.284	(0.005)	-0.363	(0.004)	-0.204	(0.004)
Mother's age at birth																
Age (divided by 10)	0.081	(0.004)	-0.355	(0.012)	-0.469	(0.005)	0.192	(0.003)	-0.282	(0.006)	0.094	(0.005)	-0.114	(0.003)	-0.107	(0.007)
Age squared (divided by 100)	0.133	(0.003)	0.303	(0.013)	0.739	(0.004)	-0.511	(0.005)	0.444	(0.008)	0.132	(0.004)	0.282	(0.004)	0.281	(0.005)
Missing	-0.150	(0.004)	0.025	(0.004)	-0.331	(0.006)	-0.553	(0.005)	-0.378	(0.005)	0.013	(0.003)	0.140	(0.003)	-0.082	(0.006)
Height at 23	0.013	(0.004)	0.600	(0.013)	0.519	(0.008)	0.419	(0.008)	-1.339	(0.004)	0.007	(0.008)	0.518	(0.006)	0.873	(0.009)
Missing	-0.385	(0.003)	0.249	(0.005)	-0.475	(0.005)	0.272	(0.005)	-1.968	(0,019)	0.134	(0.004)	-0.401	(0.003)	1.207	(0.010)
Low birth weight	0.048	(0.003)	-0.456	(0.008)	0.206	(0.003)	-0.261	(0.007)	-0.068	(0,019)	0.177	(0.003)	-0.052	(0,015)	-0.099	(0.013)
Missing	0.062	(0.008)	-0.267	(0.006)	-0.309	(0.012)	-0.344	(0.006)	-0.196	(0.004)	0.590	(0.004)	0.069	(0.004)	-0.043	(0.004)
Math score at age 7	-0.103	(0.003)	0.056	(0.003)	-0.065	(0.003)	-0.031	(0.003)	-0.527	(0.008)	0.001	(0.003)	-0.032	(0.003)	0.363	(0.007)
Missing	0.101	(0.004)	-0.955	(0,033)	0.168	(0,027)	0.038	(0.007)	0.021	(0.008)	-0.456	(0.003)	0.376	(0.003)	0.076	(0.010)
Bristol social adjustment guide at age 7	0.472	(0.005)	-0.410	(0.005)	0.144	(0.004)	0.061	(0.003)	3.269	(0,038)	-0.048	(0.003)	0.104	(0.003)	-2.046	(0,027)
Missing	-0.176	(0.011)	0.651	(0,022)	-0.085	(0,027)	-0.336	(0.005)	0.071	(0.009)	-0.100	(0.005)	-0.319	(0.004)	-0.046	(0.011)
Region of residence at birth																
Missing	-0.059	(0.004)	-0.031	(0.004)	-0.181	(0.004)	-0.331	(0.005)	-0.209	(0.004)	0.041	(0.003)	0.109	(0.003)	0.495	(0.006)
North	0.395	(0.003)	-0.119	(0.004)	-0.077	(0.003)	0.192	(0.004)	0.163	(0.003)	0.116	(0.003)	0.415	(0.004)	-0.015	(0.006)
South & Wales	0.170	(0.004)	-0.140	(0.006)	0.180	(0.003)	0.122	(0.004)	0.037	(0.005)	0.078	(0.011)	0.261	(0.005)	-0.020	(0.006)
Scotland	0.199	(0.003)	0.013	(0.006)	-0.014	(0.004)	-0.293	(0.004)	0.120	(0.004)	0.357	(0.009)	0.394	(0.006)	-0.025	(0.005)
London	0.012	(0.003)	-0.127	(0.004)	-0.130	(0.003)	-0.284	(0.005)	0.012	(0.005)	-0.347	(0.006)	0.359	(0.005)	-0.004	(0.005)
Location mass point 1	0		0		0		0		0		0		0		0	
Location mass point 2	-0.628	(0.004)	1.125	(0.004)	0.262	(0.003)	-0.884	(0.009)	-0.209	(0.004)	-0.814	(0.014)	-0.100	(0.005)	-0.529	(0.006)
Location mass point 3	-1.429	(0.008)	0.361	(0.007)	-0.546	(0.004)	-1.694	(0.005)	-1.102	(0.006)	-0.386	(0.004)	-0.634	(0.003)	-0.658	(0.014)
Location mass point 4	-2.057	(0.004)	1.486	(0.004)	-0.284	(0.004)	-2.578	(0.004)	-1.932	(0.004)	-1.200	(0.004)	-0.734	(0.004)	-1.187	(0.004)
Number of observations	12448															
Value of the -log-likelihood	105348.42															

Figure 1: Labor market states of males.

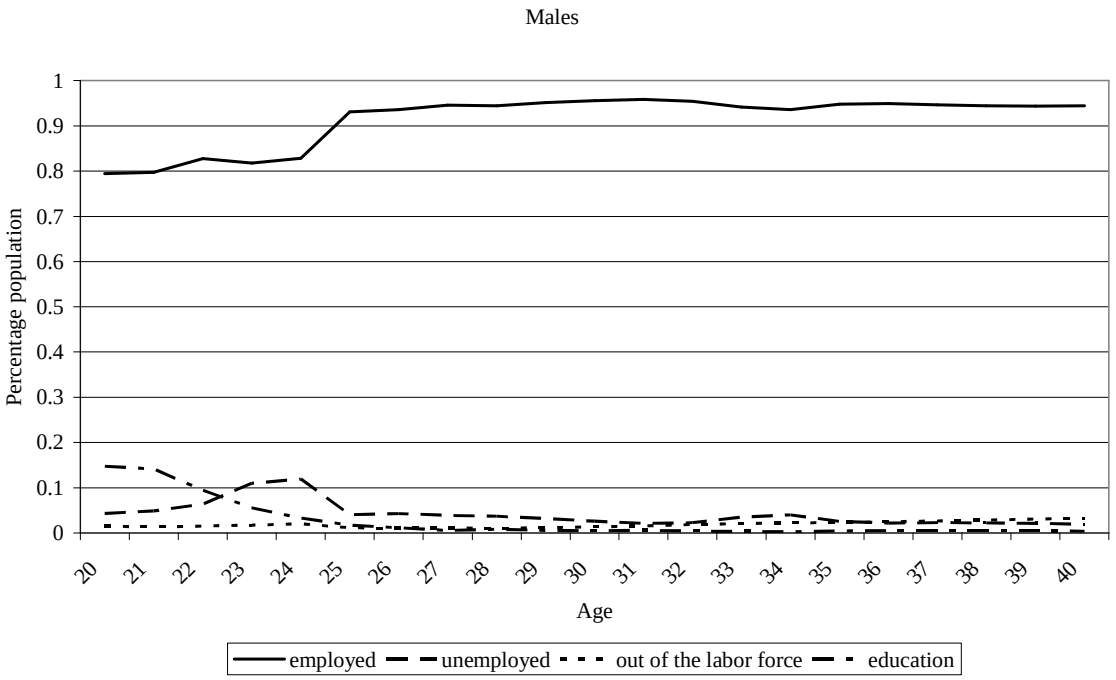


Figure 2: Labor market states of females.

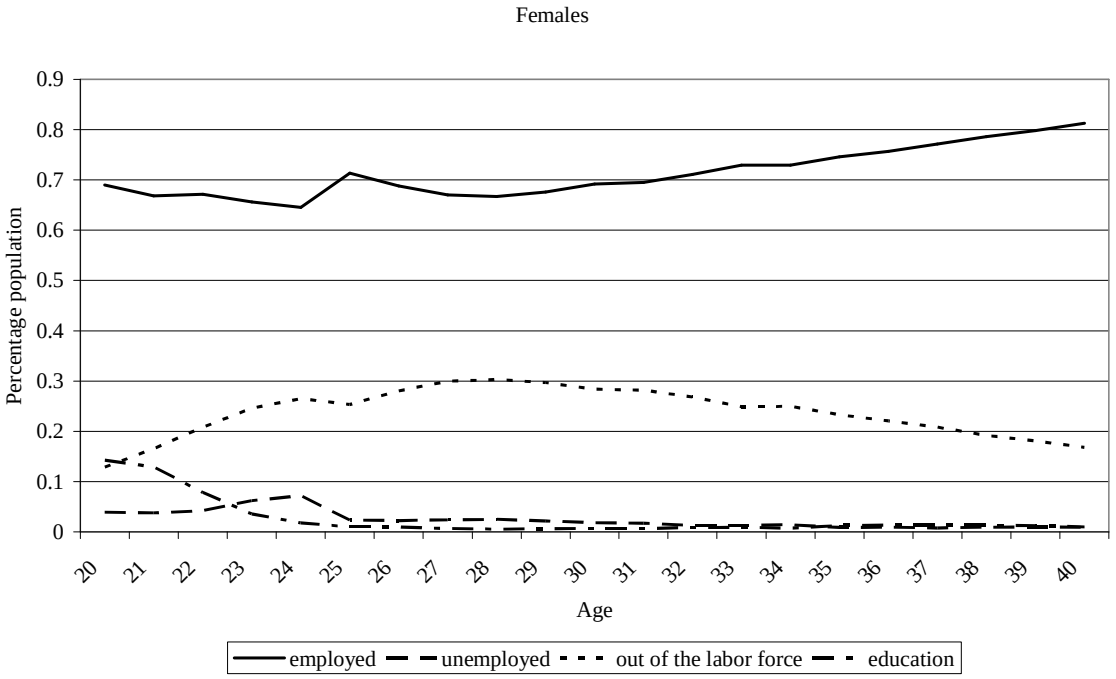


Figure 3: Disability rates of males and females.

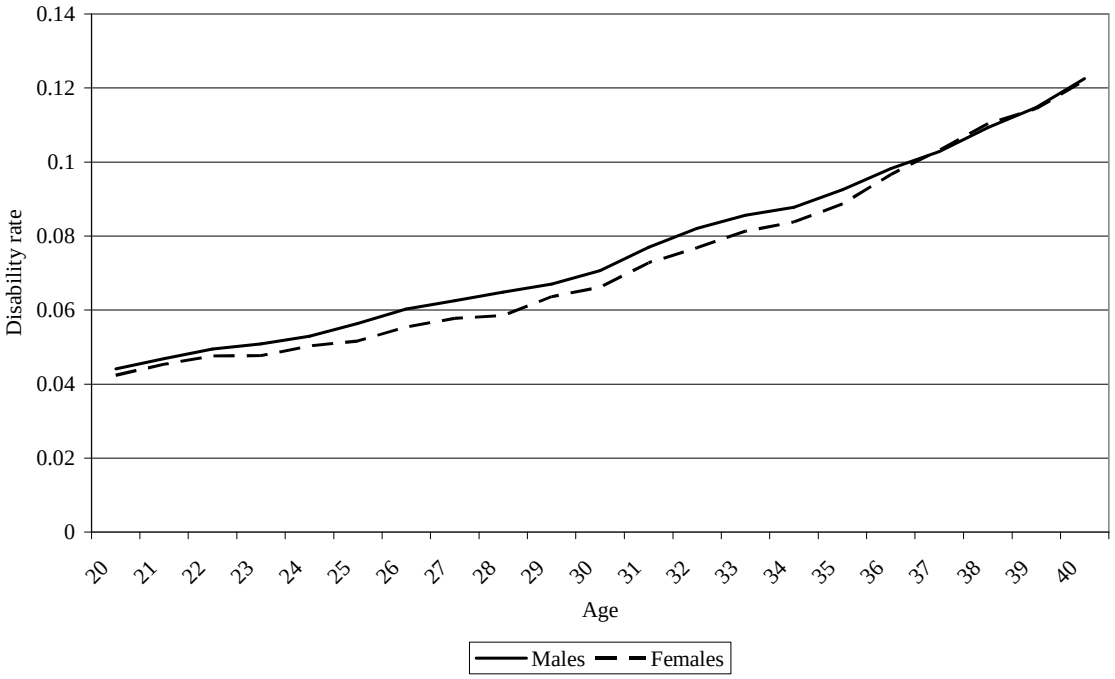


Figure 4: The annual incidence rates of health shocks for males and females.

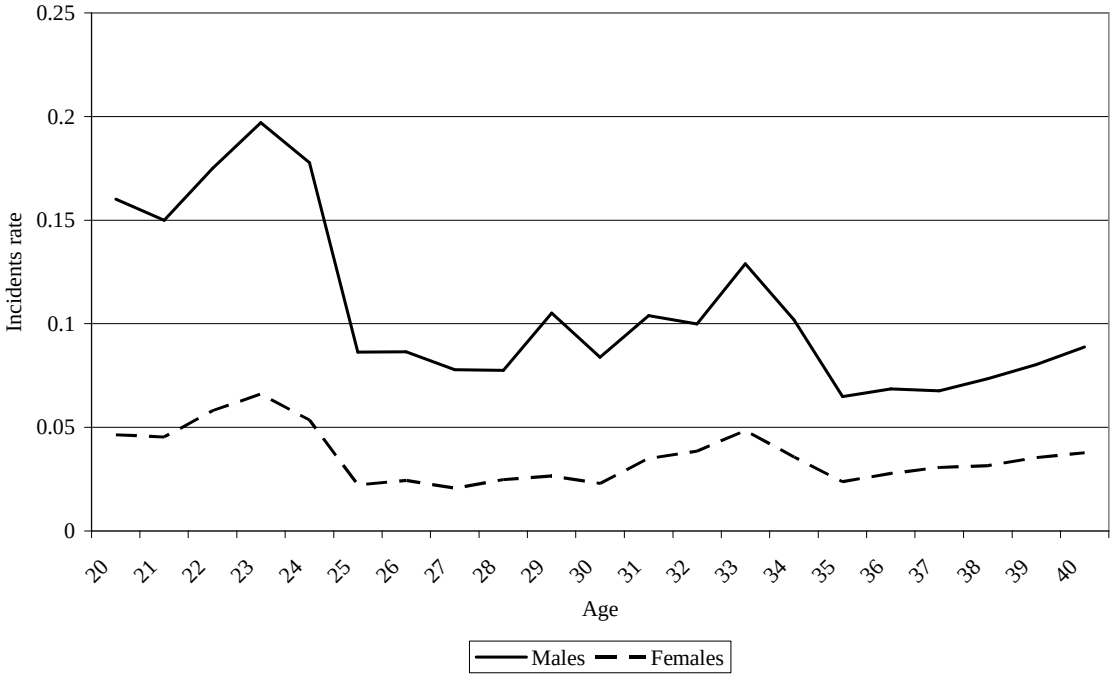


Figure 5: Disability and employment states.

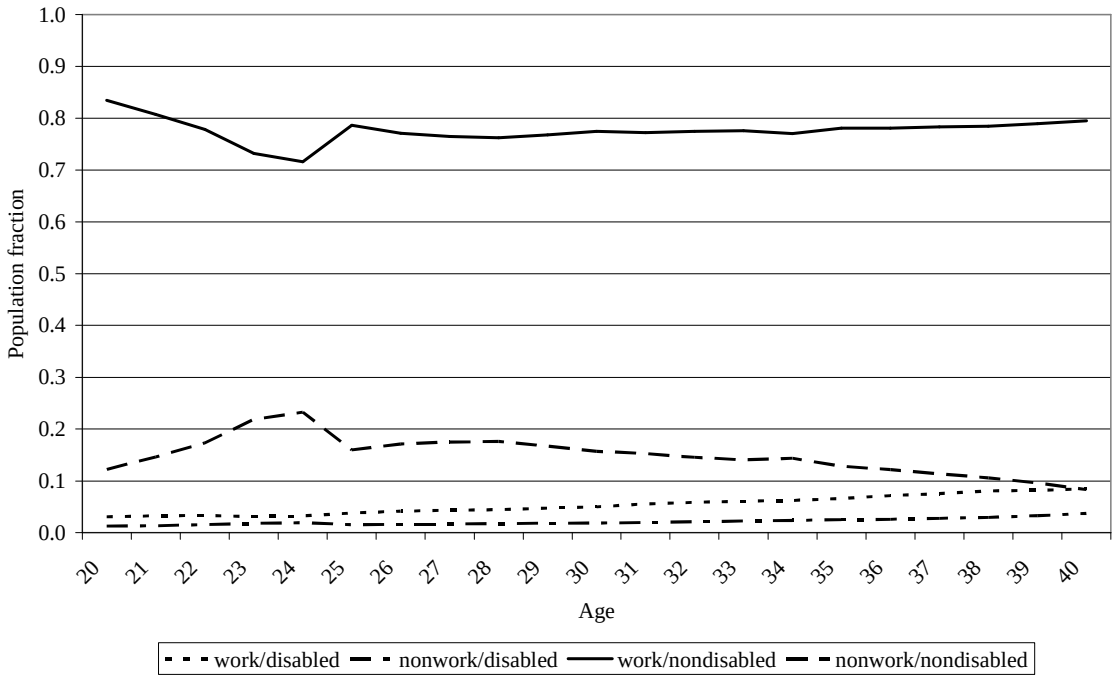


Figure 6: Effects of health shocks on the transition probabilities (coefficients)

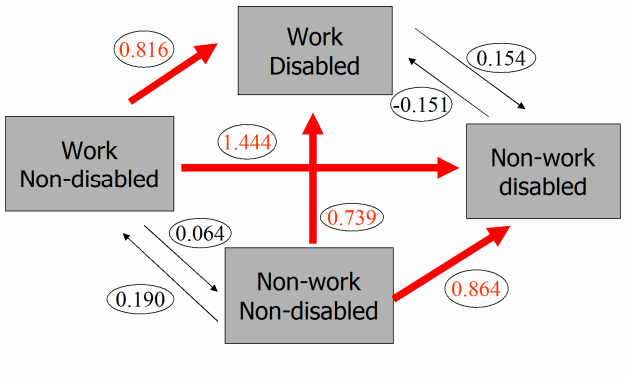


Figure 7: Employment rates for individuals with high and low socioeconomic status at birth by disability status.

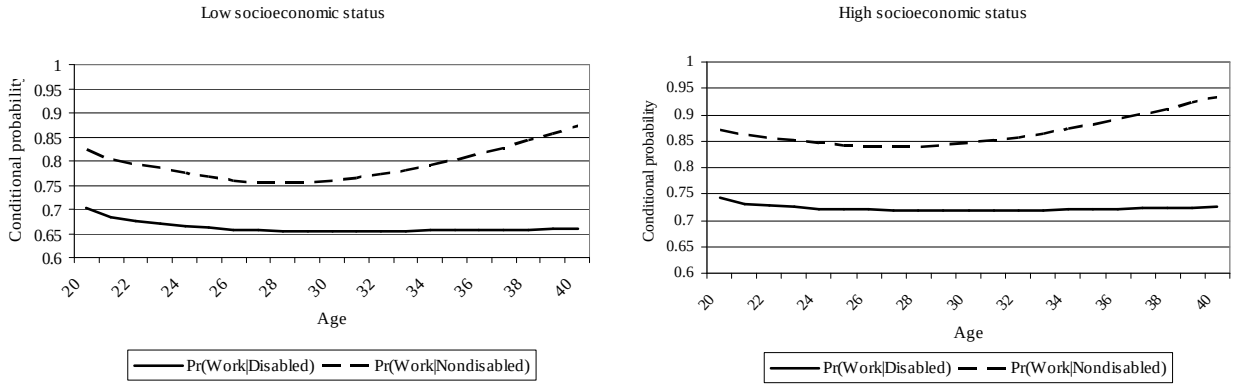


Figure 8: The effect of socioeconomic status at birth on disability rates via the initial state after school, health shocks and via transition rates: simulations with the model

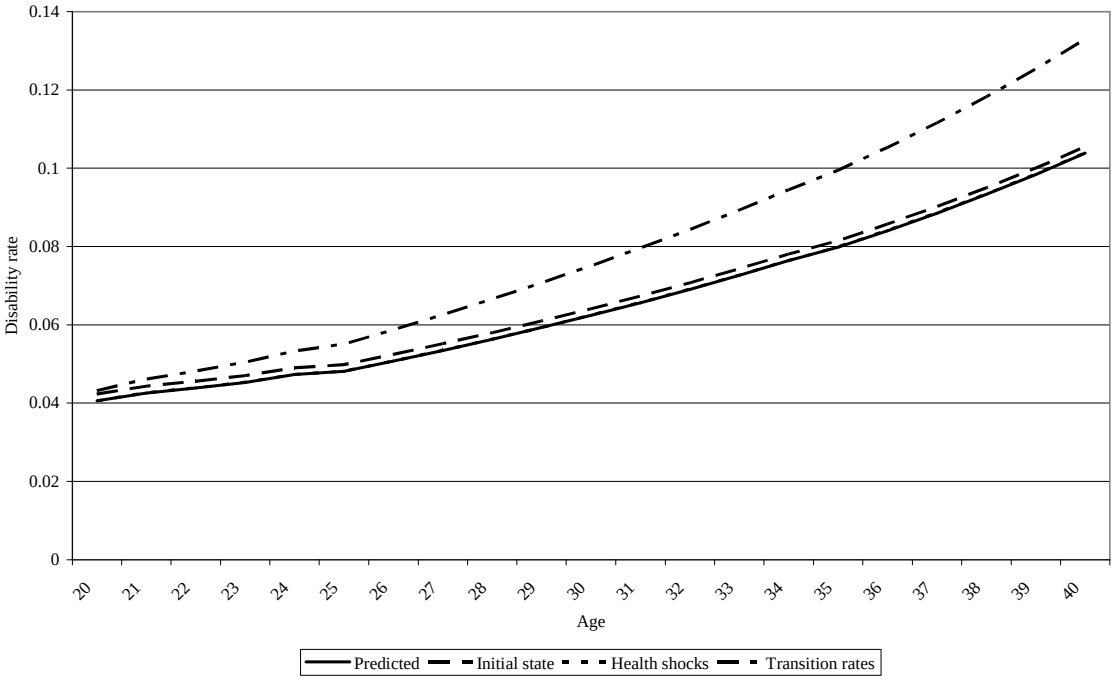


Figure 9: The effect of socioeconomic status at birth on employment rates via the initial state after school, health shocks and via transition rates: simulations with the model

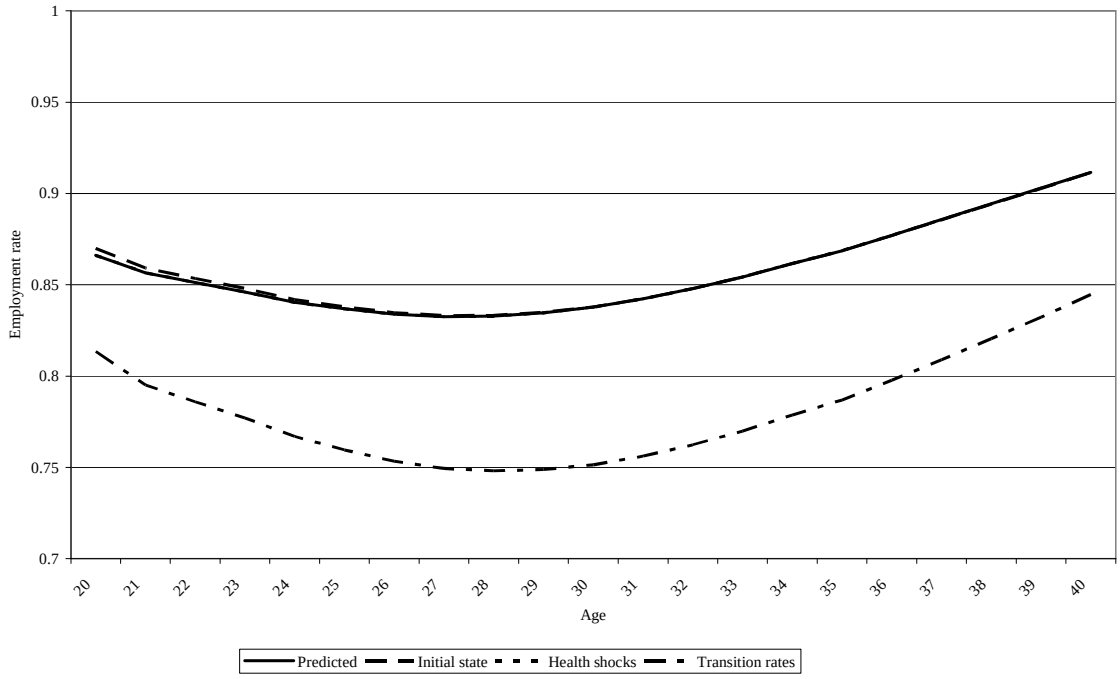


Figure 10: The effect of prenatal variables (Low birth weight and whether mother smoked during the pregnancy) on disability rates via the initial state after school, health shocks and via transition rates: simulations with the model

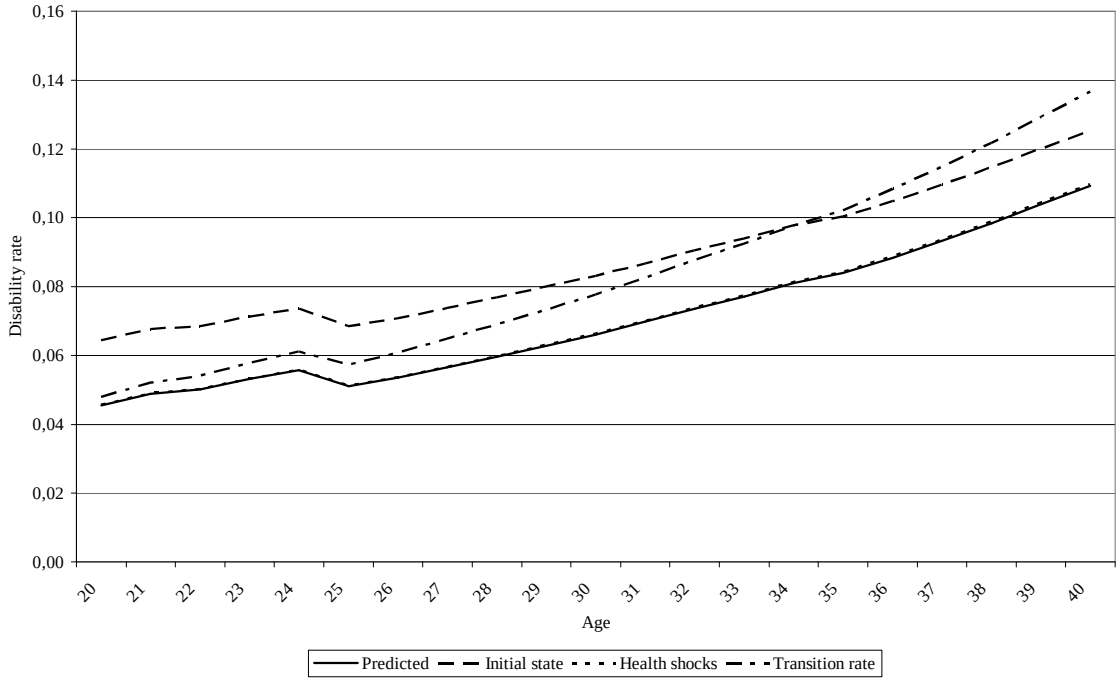
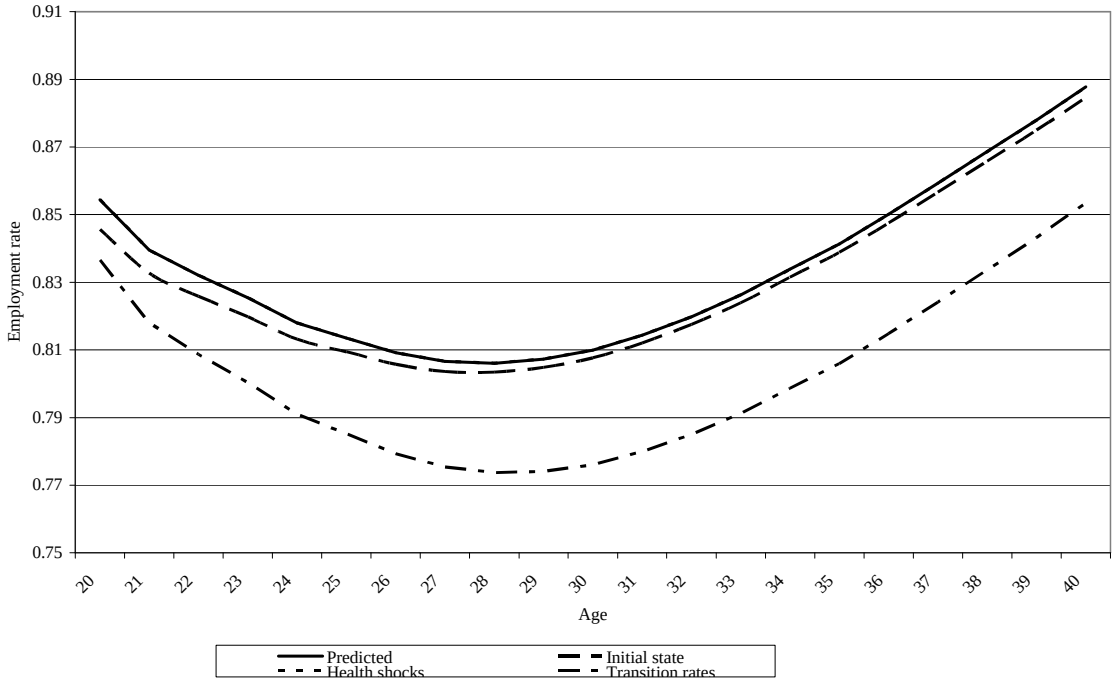


Figure 11: The effect of prenatal variables (Low birth weight and whether mother smoked during the pregnancy) on employment rates via the initial state after school, health shocks and via transition rates: simulations with the model



Appendix A: Definition of disability

We base our definition of disability on Curie and Madrian (1999) 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-9) 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).

Appendix B: Results of a multinomial model for initial state after leaving school

Table B1: Multinomial logit on the initial state

	Work / Disabled		Nonwork / Disabled		Nonwork / Nondisabled	
Intercept	0.335	(1.843)	14.409	(3.636)	-0.457	(1.128)
Gender	-0.351	(0.157)	-1.351	(0.307)	0	(0.096)
Parental socioeconomic status at birth						
Missing	0.048	(0.479)	-0.260	(1.041)	0.134	(0.301)
High	0.107	(0.136)	-0.129	(0.303)	0.376	(0.080)
Low	0.193	(0.130)	-0.343	(0.280)	0.209	(0.084)
Mother's smoking at pregnancy	0.135	(0.114)	-0.013	(0.237)	0.154	(0.070)
Missing	-0.671	(0.648)	-0.17	(1.030)	0.213	(0.276)
Mother's age at birth						
Age (divided by 10)	0.131	(0.737)	-0.713	(1.594)	-0.312	(0.493)
Age squared (divided by 100)	-0.066	(1.263)	1.920	(2.630)	0.895	(0.834)
Missing	-21.074	(3.047)	-8.129	(1.755)	-10.135	(1.102)
Height at 23	-1.962	(0.792)	-9.827	(1.554)	-1.429	(0.481)
Missing	-2.482	(1.409)	-15.868	(2.714)	-2.396	(0.902)
Low birth weight	0.396	(0.205)	0.550	(0.346)	0.172	(0.141)
Missing	-13.735	(1.764)	1.504	(1.003)	-0.236	(0.594)
Math score at age 7	-89.830	(25.459)	-284.888	(59.619)	-13.059	(15.389)
Missing	-0.187	(0.465)	-2.264	(0.704)	0.384	(0.273)
Bristol social adjustment guide at age 7	16.479	(6.220)	36.636	(12.157)	20.903	(4.040)
Missing	-0.169	(0.462)	1.922	(0.663)	-0.021	(0.264)
Region of residence at birth						
Missing	35.601	(3.610)	6.170	(1.909)	10.698	(1.115)
North	-0.159	(0.143)	0.418	(0.322)	0.344	(0.092)
South/Wales	-0.108	(0.161)	0.731	(0.343)	0.273	(0.105)
Scotland	-0.261	(0.196)	-0.042	(0.448)	0.195	(0.122)
London	-0.373	(0.173)	-0.268	(0.43)	0.035	(0.109)
Number of observations			12448			
Value of the - Log Likelihood function			5273.17			