CenEA Working Paper Series
WP07/09

Dynamics of poor health and non-employment

Peter Haan
Michał Myck

Abstract:
While there is little doubt that the probability of poor health increases with age, and that less healthy people face a more difficult situation on the labour market, the precise relationship between facing the risks of health deterioration and labour market instability is not well understood. Using twelve years of data from the German Socio-Economic Panel we study the nature of the relationship between poor health and non-employment on a sample of German men aged 30-59. We propose to model poor health and non-employment as interrelated risks determined within a dynamic structure conditional on a set of individual characteristics. Applying dynamic panel estimation we identify the mechanism through which poor health contributes to the probability of being jobless and *vice versa*. We find an important role of unobserved heterogeneity and evidence for correlation in the unobservable characteristics determining the two processes. The results also show strong persistence in the dynamics of poor health and non-employment.

Centre for Economic Analysis
CenEA
www.cenea.org.pl

Dynamics of poor health and non-employment*

Peter Haan[†] and Michal Myck[‡]

April 29, 2009

Abstract

While there is little doubt that the probability of poor health increases with age, and that less healthy people face a more difficult situation on the labour market, the precise relationship between facing the risks of health deterioration and labour market instability is not well understood. Using twelve years of data from the German Socio-Economic Panel we study the nature of the relationship between poor health and non-employment on a sample of German men aged 30-59. We propose to model poor health and non-employment as interrelated risks determined within a dynamic structure conditional on a set of individual characteristics. Applying dynamic panel estimation we identify the mechanism through which poor health contributes to the probability of being jobless and *vice versa*. We find an important role of unobserved heterogeneity and evidence for correlation in the unobservable characteristics determining the two processes. The results also show strong persistence in the dynamics of poor health and non-employment.

Keywords: Health, non-employment, risk, ageing, dynamic panel data.
JEL Classification: C33, J21, J14.

*The authors thank Viktor Steiner, Kostas Tatsiramos and Arne Uhlendorff for valuable comments. We are also grateful to seminar participants at the Geary Institute and DIW-Berlin for comments and suggestions. Financial support from the German Science Foundation (DFG) through project 'Economic risks, savings of the elderly, and the impact of public policy - an empirical analysis for Germany" (STE 681/8-1) is gratefully acknowledged. The usual disclaimer applies.

[†]DIW Berlin, phaan@diw.de
[‡]DIW Berlin, CenEA and IFS; mmyck@diw.de

Introduction

We examine the nature of the relationship between poor health and non-employment. It is rather well established that poor health is one of the key determinants of labour market inactivity, and an important factor driving individuals out of work and reducing their probability of entry into employment (e.g. Chan and Stevens (2001), Kalwij and Vermeulen (2008)). This is especially the case for older workers as health deteriorates with age. In many cases in the literature, however, the relationship between health and work is considered as a one-way causality with health being the determining factor which affects choices on the labour market (e.g. Blundell et al. (2002)). The interpretation in this case is that poor health is the principal risk individuals face and employment (or even labour market participation) is the choice made conditional on health status.

The relationship between non-employment and poor health is, however, likely to be more complex. Unhealthy individuals may *choose* not to work (or even no to look for work) but it is also likely that poor health may reduce productivity and thus increase the risk of job loss. Similarly, it may also make finding new employment more difficult. At the same time one can also imagine a possible effect of lack of employment on health, with long spells out of work having detrimental consequences for both mental and physical well-being (e.g. Clark and Oswald (1994)). This in turn may further worsen individuals' prospects on the job market. Chan and Stevens (2001) and Kuhn and Sweetman (1999) show that the effects of job separations on future employment prospects and re-employment are very strongly related to age. Such findings are typically associated with high reservation wages and loss of job-specific human capital, but deteriorating health may be an important factor behind them. Because of a strong relationship of both health and employment with age and because of their likely correlation it seems natural to model the two processes together. Endogeneity of health and employment has been recognised in the literature for some time (e.g. Brown et al. (2005)), but to our knowledge our approach is the first attempt to take advantage of the panel dimension of the data and to model health and employment processes jointly while accounting for unobserved individual effects influencing both processes.

A clear understanding of the relationship between poor health and non-employment should also shed light on the role of ageing in the determination of the labour market status. We know that with age the likelihood of poor health increases and the probability of working falls, but the role of health in the latter process, and the precise relationship between employment and health is not well understood. From the point of view of the policymaker it is important to separate out the effect of deteriorating health from the effect of age *per se* as a factor increasing both risks. Better understanding of the relationship between these two processes should help in addressing such issues as age discrimination in the labour market and the role of health improvements in extending labour market activity. Health is likely to be an important factor in determining employment, but given the complex nature of the relationship between the two outcomes it is unclear how improvements in one of them would be reflected on the other.[1]

In the paper we develop a joint intertemporal model of health and non-employment which is estimated on a sample of men aged 30-59 using the German Socio-Economic Panel data (SOEP) for the years 1996-2007. We model the two processes jointly in a discrete sequential equation model where we assume that the health risk affects the non-employment risk and vice versa. We estimate a dynamic bivariate logit model in which we explicitly account for the joint distribution of unobserved heterogeneity in a non-parametric way and we control for the initial conditions as in Wooldridge (2005). Similar to Alessie et al. (2004) we assume sequential causality, in which the last period's health is assumed to affect the current period's employment and the last period's employment is assumed to affect this period's health.

The results confirm a strong and significant relationship between the risks of poor health and non-employment, and underline the separate important role of ageing in determining the two outcomes. We find evidence for positive correlation in unobservable characteristics determining the risk of poor health and non-employment which indicates that a separate treatment would lead to an overestimation of the relationship. Still we find *ceteris paribus* a positive effect of bad health on the employment risk (and *vice-versa*). Because of the nature of the dynamics of the two processes this effect increases with age. Our findings show the importance of health status for employment outcomes all the way throughout the life course, but particularly after the age of about 50. The estimated dynamics of the relationship suggest a particularly important role of lack of employment early in the life-cycle for employment outcomes after the age of about 45. From the policy perspective our results confirm the role played by employment in the early stages of life for the development of labour market activity and suggest a significant role of health-related policies throughout the working lives for increases in the level of employment particularly in the years close to retirement.

We begin the paper by describing the data we use for the estimation and by specifying the

[1]From the policy perspective the interaction of health and employment will also be important from the point of view of costs of poor health and the return to spending on health improvements (an example of trying to put a value on health improvements can be found in Manton (1998)).

details of the risks we model (Section 1). We then present the econometric specification in Section 2, which is followed by results and conclusions (respectively Sections 3 and 4).

1 Data

The data we use in this paper are taken from the German Socio-Economic Panel (SOEP) and we use the waves covering the period from 1996 to 2007. The SOEP is a representative sample of all private households in Germany and annually collects information on over 11,000 households. In our application we focus on men aged 30-59 and the estimation is conducted using an unbalanced sample of individuals observed for at least three consecutive waves. This restriction results from the requirement of having at least two consecutive waves for the identification of unobserved heterogeneity and an additional one for information on the first lag of the dependent variables. The focus on men and the age restrictions relate primarily to the interpretation of non-employment as a 'risk', i.e. to the assumption concerning the involuntary nature of non-employment, which in the case of women and of men below 30 or above 59 would seem stronger and would be more difficult to justify.[2]

1.1 Poor health and non-employment

Our analysis covers two types of outcomes, poor health and non-employment. These outcomes are modelled as stochastic dynamic processes conditional on several individual and household characteristics which affect the probability of observing these outcomes. The categorisation of people into those experiencing the realisation of the two types of risk is done on the basis of information on their health status assessment and employment available in the SOEP data we use in the estimation.

In all waves of data the individuals are asked to rate their health status at the day of the interview by choosing one of five options: very good, good, satisfactory, poor, very poor. In our application we classify people as being in "poor health" when they declare themselves into either of the last two categories. While such self-reported poor health is clearly a subjective measure of health status, it is likely to encompass various aspects of health which may affect the situation on the labour market, including both physical and mental well-being. Self-reported health measures are commonly used in the literature and have been often demonstrated to be good overall reflections of objective health status and valid longitudinal measures of changes in health (e.g. Benitez-Silva and Ni (2008)). Using an international comparison based on the

[2]As a robustness check we further restrict the sample to men aged 30-54.

Survey of Health, Ageing and Retirement in Europe (SHARE) Kalwij and Vermeulen (2008) showed that in Germany objective health measures add little on top of self-reported health information in a cross-sectional analysis of employment patterns.[3]

The non-employed are identified through the survey information on weekly working hours prior to the interview date. We define all individuals as non-employed who report to work zero working hours in their 'actual working week'.

It is not surprising that both of the risks we analyse depend very strongly on age. As we can see in Figure 1A the employment rate for men in Germany in the period covered by the data falls from 93.4% for men aged 37 to 73.9% for those aged 55 and 47.1% for those aged 60. Interestingly at the same time the ILO unemployment rate (Figure 1B) changes from 3.4% to 7.6% to 2.9% of the male labour force. This suggests significant movements out of the labour force into inactivity. It is likely that health-related factors play an important role in this process. Figures 1A and 1B also show the difference in labour market outcomes by education level. Those with more than 12 years of schooling are classified as "education 1", while those with 12 years or less as "education 2". Employment levels are significantly lower among those in the lower education group and the difference grows with age. For example employment among men aged 55 in the higher education group is 87.6% while among the lower educated only 68.0%. Unemployment levels are also much higher among the lower educated but in this case the difference between education groups does not seem to be as strongly related to age as in the case of employment.

When we look at the development of health assessment (Figure 1C), poor health among men remains below 10% up until the age of about 40 but then health deteriorates relatively quickly. 17.0% of men aged 50 declare poor or very poor health and this proportion grows to 22.9% and peaks at the age of 57 (25.4%). There is a divergence in the proportion of men in poor health by education level from the age of about 38. While for those aged 38 the difference is only 2.6 percentage points, it grows to 7.8 percentage points for those aged 50 and to 9.2 percentage points among those aged 55.

[3] The SOEP contains also information on the official disability status. There are several reasons why using this information is not well-suited for our purpose. Primarily the official disability status is a rather narrow definition of poor health and may neglect important features of health, especially from the point of view of labour market activity. Many medical conditions would not be classified as "disability" and yet would have profound consequences on the ability to work. What also sheds some doubt about the validity of using this variable is the fact that only 54% of those with disability levels above 50% (i.e. those who in the system are classified as severely disabled) assess their health status as lower than satisfactory. Similarly only 55% of those with the disability status declare that their health hinders the fulfilment of their daily activities. In the case of those who declare poor or very poor health status this proportion is 83%. Using the self-reported health status also means that we avoid the possible endogeneity of disability resulting from disability insurance incentives (see, e.g. Gruber (2000), Riphahn (1999)), though evidence from the 1980s may suggest that in Germany this problem is not as serious as in some other countries (see Bratberg (1999)).

In Figure 1D we demonstrate how strongly employment varies by the declared health status. The differences are shown conditional on education group, and as we can see are especially large among those in the lower education group. Among those aged 35 the difference in employment rates between those in good health and those in poor health is only 2.8 for the higher education group, and as much as 19.1 percentage points for the lower educated. Among the latter group the difference is as high as 28.4 percentage points for individuals aged 48, while for the better educated it grows to over 20 percentage points among those aged 54.

Such strong dependence on age raises interesting issues related to the dynamics and the relationship between health and employment and poses questions about the importance of ageing versus the importance of health deterioration for labour market outcomes. With a long panel data at our disposal we can address both of these problems and try to identify the separate role of health and age in determining non-employment, as well as the role of lack of labour market activity on health.

1.2 The estimation sample

For the purpose of our estimation we limit the sample to those aged 30-59 with at least two consecutive observations in the panel (t, t+1, t+2, ...) and one lagged observation (t-1). The latter is necessary because of the dynamic nature of the specification. After dropping observations with missing crucial information as well as excluding the civil servants we end up with a sample of 4,420 individuals observed for at least two consecutive waves (plus the initial lagged observation) and a total sample of 26,999 observations, which gives us the average of 6.1 observations per individual. In Table 1 we present some descriptive statistics on this estimation sample. The average initial age, i.e. the age at the first observation, is 41.4 while the average age at the last observation is 46.5. About 18% of the sample are single individuals, and the majority (about 74%) live in the states which constituted the former West Germany. The average number of years of continuous full-time education is 12.6.

In the first wave of observation 11.3% of individuals declare poor or very poor health status, and this proportion grows to 15.9% for the last observation. The proportions of non-employed are respectively 13.8% and 19.3%.

Figure 2 shows the relationship between age and the dependent variables in our estimation sample. The proportion of those in poor health increases from about 5% in the early 30s to 17.9% for those aged 50, to 24.9% for those aged 55, and falls slightly to 23.8% for individuals aged 59 (Figure 2A). Non-employment oscillates between 8-11% up to the age of 45 and then grows to 15.7% for those aged 50, 25.7% for those aged 55 and reaches 47.2% for those aged 59. Thus while health status seems to level out at the age of about 55, non-employment grows significantly especially in the late 50s. While some of this growth is clearly related to factors other than health (e.g. early retirement or increasing involuntary job separations), those in poor health are still much more likely to be non-employed (see Figure 2B) compared to those in good health. Moreover, among healthy individuals, only a small fraction of the increase in non-employment is due to pension receipt. As we can see in Figure 2C the proportion of non-employed individuals also grows among those in good health who are not receiving pensions from about 10% among those aged 50 to 33.2% among those aged 59. We examine the importance of this divergence of poor health and non-employment rates by estimating our model on the full sample of men aged 30-59 and also on a restricted sample aged 30-54. As we shall see overall the results of the two estimations are very similar.

2 Econometric specification

To analyse the intertemporal effects of poor health on non-employment we derive and estimate a joint discrete intertemporal model of health and employment risks.

In general, intertemporal models based on panel data suffer from the problem of initial conditions. This problem arises as the health and employment status of an individual observed in the initial observation $t = 0$ cannot be assumed to be random. In fact it is likely that non-random unobservable effects are correlated with the status in the initial state. To account for initials conditions we follow Wooldridge (2005) and model the distribution of the individual specific random effects conditional on information of the initial observation. This method is operationalised by allowing the mean of the random effect appearing in the health and employment processes to depend on the first observation of the individual's employment and health status. Moreover, we condition the random effects on working experience prior to the initial observation, expressed as the ratio of experience to age at the first observation.[4] . The remaining stochastic element of the individual random effect, denoted with c_i^u and c_i^h, is specified below.[5]

In period t=*1,...,T* each individual, indexed i, has a positive probability of making transitions between the 'risk free' employment state and non-employment. We denote non-employment by

[4]This treatment accounts for the age heterogeneity of the sample at the first observation

[5]Akay (2009) conducts several Mote Carlo simulation to study the performance of the Wooldridge method. He finds that in moderately long panels with more than 5 periods the methods has a comparable good performance relative to other procedures that account for the non-random selection.

Table 1: Sample statistics, estimation sample

Number of individuals:	4420			
Number of observations:	26,999			
	Mean	Std. dev.	Min	Max
Age:				
- first observation	41.396	(9.15)	30	58
- last observation	46.505	(8.83)	31	59
Experience:				
- first observation	20.311	(9.85)	0	45
- last observation	24.947	(9.54)	0	46
Lives in west				
- first observation	0.737	(0.44)	0	1
- last observation	0.738	(0.44)	0	1
Has children:				
- first observation	0.610	(0.49)	0	1
- last observation	0.562	(0.50)	0	1
Number of children (conditional)				
- first observation	1.734	(0.80)	1	8
- last observation	1.718	(0.79)	1	8
Single:				
- first observation	0.184	(0.39)	0	1
- last observation	0.182	(0.39)	0	1
Self-reported poor health:				
- first observation	0.113	(0.32)	0	1
- last observation	0.159	(0.37)	0	1
Non-employment:				
- first observation	0.138	(0.34)	0	1
- last observation	0.193	(0.39)	0	1
Number of waves per individual	6.108	(2.89)	2	11
Years of education	12.662	(2.74)	7	18
County-level unemployment	12.32	(5.49)	2.81	32.1

Source: GSOEP data, waves 1997 - 2007.

u and the individual's probability of being in this state at time t is given by P_{it}^{u}.

There are numerous channels through which individuals might become non-employed. These include demand side factors, such a technological change or economic shocks, as well as supply considerations. Both of these sets of factors are likely to be affected by age, education, the socio-economic background and the lagged labour market status. In this analysis we focus in particular an an additional source of being at risk on the labour market, the risk of being in poor health. In our approach all non-employment is considered as realisation of the labour market risk. The main reason for this assumption is the fact that any other definition of labour market risk, for example limiting it to ILO unemployment, is likely to be affected by health status. Unhealthy people will be less likely to look for work, but if their non-employment results form poor health then it is difficult to argue that they are not "at risk" on the labour market. In contrast to other demographic characteristics we model the health status endogenously as a dynamic latent process which describes the individual probability P_{it}^{h} of facing the health risk. We allow for mutual interaction between health and employment risks and condition the health process on non-employment.

Similar to Alessie et al. (2004) we specify the non-employment and health risks jointly in a bivariate dynamic random effects model. The key advantage of the joint modelling is the possibility to account for unobservable characteristics which might jointly affect both processes. In this way it is possible to separate the direct interaction between the health and employment states from indirect unobserved sources.

Amongst others Carrasco (2001) discusses identification problems in the so called "dummy endogenous variable models". The key problem is the simultaneity of the two latent processes and in order to identify causal relationships exclusion restrictions are required. Carrasco (2001) and similarly Michaud and Tatsiramos (2008) estimate a joint dynamic model of fertility and female employment and use the sex composition of children as an instrument for fertility in the employment equation. In the model proposed above at least two different exclusion restrictions are required as both endogenous variables can be assumed to simultaneously affect the other process.[6]

We have experimented with several sources of identification. In our view the most promising exclusion restriction for the health process are parental health information (specifically infor-

[6] A detailed discussion about the identification of causal effects of socio-economic variables and health can be found in a special issue of the Journal of Econometrics 112 (2003). In this issue Adam et. al. (2003) mention the lack of appropriate exclusion restrictions and propose instead to use Granger causality tests to rule out causality between health status and socio-economic variables.

mation if the father is still alive and if not at what age he died) which should affect the health status of the child but not his/her labour market risk. Moreover, risky behaviour such as years of smoking should in principle affect the current health status but not the non-employment risk. A potential exclusion restriction for non-employment is the regional unemployment rate. Brown et al. (2005) argue similarly and provide evidence that parental information serves as a good instrument in their application based on cross sectional data. Using only the cross sectional variation our results support their findings but once accounting for the individual effects in the panel dimension the explanatory power of the instrument becomes very weak. This suggest that the parental information is rather a proxy for the individual unobservables and should not be used as an instrument for the current health status.

Therefore, in line with Alessie et al. (2004) we specify a sequential intertemporal latent model for health and non-employment risk. In particular we assume that the non-employment in period t is affected by the last period's health status and similarly the risk of poor health is affected by the labour market status in the period $t-1$. Since we control for the initial conditions problem, jointly account for unobservables in the two processes and allow those effects to be correlated, the sequential timing circumvents the simultaneity problem and provides identification of the intertemporal effects. This identification strategy rules out that the stochastic shocks ϵ_{1it} and ϵ_{2it} affecting both processes are correlated over time.

This approach is also justified by the nature of the data where we observe both risks at the time of the interview. This makes it impossible to determine the precise timing of the two processes. In the sequential model we specify the timing of the relationship, though due to the nature of the data the periodicity cannot be shorter that one year. However, given the weakness of the aforemention exclusion restrictions and the missing information about the timing we think that the sequential treatment of the two risks seems to be a clearer and less controversial treatment than using a true simultaneous model.

Our intertemporal model takes the following form:

$$h_{it}^* = h_{it-1}\gamma_h + u_{it-1}\alpha_h + h_{i0}\delta_h + x_{1it}\beta_h + c_i^h + \epsilon_{1it} = X_{1it}\beta_1 + c_i^h + \epsilon_{1it} \quad \text{for} \quad t = 1, ..., T. \quad (1)$$

$$u_{it}^* = u_{it-1}\gamma_u + h_{it-1}\alpha_u + u_{i0}\delta_u + x_{2it}\beta_u + c_i^u + \epsilon_{2it} = X_{2it}\beta_2 + c_i^u + \epsilon_{2it} \quad \text{for} \quad t = 1, ..., T. \quad (2)$$

The latent model of being in poor health is assumed to depend on observable characteristics x_{1it}, including the lagged health status h_{it-1} and the health status in the initial state h_{i0}, and on an indicator for non-employment in the last period u_{it-1}. Moreover, the health status is affected by c_i^h which is an individual unobserved effect, and by ϵ_{1it} which is an iid error we assume to follow a logistic distribution.

The risk of non-employment depends on lagged non-employment u_{it-1}, the initial state u_{i0}, observable characteristics including the regional labour market situation x_{2it}, an individual specific unobserved effect c_i^u and an error term ϵ_{2it} which we assume again to be iid and to follow a logistic distribution. In addition the latent model of non-employment is affected by health status in the last period h_{it-1}. An individual is in poor health ($h_{it} = 1$) or is non-employed ($u_{it} = 1$) if the respective latent variable exceeds some threshold c^*.

In contrast to Alessie et al. (2004) who estimate the dynamic bivariate model with a parametric bivariate probit model, we assume that the individual specific random terms c_i^u and c_i^h are jointly non-parametrically distributed and follow a discrete mass point distribution as in Heckman and Singer (1984). Conditional on the individual specific unobserved effects and assuming that ϵ_{1it} and ϵ_{2it} follow a logistic distribution, the joint probabilities of health and employment risks are described by the following four combinations (for clarity of presentation we leave out the indexes for individuals and time):[7]

$$P(h = 1, u = 1|c^h, c^u) = \frac{\exp(X_1\beta_1)}{1 + \exp(X_1\beta_1)} \times \frac{\exp(X_2\beta_2)}{1 + \exp(X_2\beta_2)} \tag{3}$$

$$P(h = 1, u = 0|c^h, c^u) = \frac{\exp(X_1\beta_1)}{1 + \exp(X_1\beta_1)} \times \frac{1}{1 + \exp(X_2\beta_2)} \tag{4}$$

$$P(h = 0, u = 1|c^h, c^u) = \frac{1}{1 + \exp(X_1\beta_1)} \times \frac{\exp(X_2\beta_2)}{1 + \exp(X_2\beta_2)} \tag{5}$$

$$P(h = 0, u = 0|c^h, c^u) = \frac{1}{1 + \exp(X_1\beta_1)} \times \frac{1}{1 + \exp(X_2\beta_2)} \tag{6}$$

We specify the unobserved individual terms c_i^u and c_i^h in a joint non-parametric distribution. As is common in dynamic nonlinear models, we specify the individual specific unobservables as random effects. The random effects are described by three points of support (mass points) for the latent health model (c_1^h, c_2^h, c_3^h) and the propensity of non-employment (c_1^u, c_2^u, c_3^u) which load on one probability vector. This specification is flexible enough to freely estimate the correlation of the unobservables underlying the two processes.

[7] Note that this specification implies that only the individual specific unobservables affecting the health and employment risk are correlated. The time specific shocks are assumed to be independent.

Mass points and the transformed probabilities π are jointly estimated with the parameters by maximum likelihood.[8] The likelihood to be maximised is then:

$$L = \prod_{i=1}^{n} \sum_{k=1}^{3} \pi_k \prod_{t=1}^{T} \prod_{j=1}^{J} Pr(P_{it} = j)^{d_{itj}}, \tag{7}$$

where $d_{itj} = 1$ if j is the observed probability combination and 0 otherwise.

3 Results

In Table 2 we present results of the specification outlined above (Specification 2) together with a separate estimation of the two processes which does not account for unobserved heterogeneity (Specification 1). As far as the exogenous explanatory variables are concerned, due to lack of significance and stability in the estimated values of some variables we have excluded marital status, nationality and regional unemployment variables from the health risk equation.

Most importantly we find a significant interaction between the health and the employment process. As expected this interaction can be partly related to unobserved individual specific effects and to the effect of the lagged employment and lagged health status. Specification 2 shows a positive and significant correlation of about 0.5 in the unobservables driving the two processes, and this explains why the estimated coefficients of the lagged employment on the current health status and of the lagged health status on the current employment status are lower than in Specification 1.[9] This implies that a separate specification without unobserved heterogeneity leads to upward biased effects of the employment and health variables. Moreover, the significance of both variables in Specification 2 underlines that it is necessary to allow for the mutual interaction when modelling the employment and health processes over time even when controlling for the individual effects. Poor health at time *(t-1)* increases the probability of non-employment, and there is evidence that lack of employment at *(t-1)* increases the risk of poor health at time t. These findings are in line with previous literature which has separately shown that lack of employment has a detrimental effect on physical and mental health (e.g. Clark and Oswald (1994)) and that bad health leads to lower employment. In Section 3.1 we present several simulations to show the magnitude of the estimated effect of the health risk on the risk of non-employment.

Further, in line with the previous literature, the results show a very strong level of persistence

[8] Probabilities are transformed to guarantee positive probabilities that add up to unity.

[9] According to the Akaike-Criterium which is based on the differences in the log likelihood, the unobservables significantly affect the two process.

in both processes which is reduced when unobservable characteristics are controlled for. This state dependence in non-employment can be related to higher search costs, stigma effects or human capital depreciation of the unemployed, see e.g. Hyslop (1999). The persistence in the health status indicates a long term nature of health conditions and may reflect long-term illnesses and chronic diseases.

The coefficients on exogenous explanatory variables all have the expected signs. Education and number od children reduce both the health and the employment risk. The age terms have the expected significant effect on both poor health and non-employment, and the risk of non-employment is *ceteris paribus* higher for singles and foreigners. The regional unemployment rate has the expected strong effect on the probability of non-employment.

As a robustness check given the divergence of the trends of the two risks after the age of 54 (see Figure 2A) we also present results on an age-limited sample of individuals aged between 30-54 to show the stability of our findings. These results are shown in Table 3. The sign of all coefficient remains the same and their magnitude changes significantly only in the case of the age effects in the non-employment equation. This is not surprising given the pattern we see in Figure 2A. Most importantly the coefficient of the lagged health and the lagged employment status are very similar to coefficients estimated on the full sample.

3.1 Simulation of employment risk by health status

In order to get a better understanding of the magnitude of the estimated relationship between poor health on non-employment we perform several simulations of a stylised individual. In particular we simulate the probability of employment over the life-cycle under different health/employment scenarios. The simulations differ either by initial conditions or by the timing of the onset of poor health. All other characteristics of the modelled individual including his unobserved effects remain constant. The differences in the development of the employment rate show the effect of different initial conditions (Figure 3) and of the realisation of the risk of poor health at a specific age (Figure 4).

We simulate employment rates of a married man of German nationality with two children, average eduction of 12 years, working experience at the initial state (at the age of 29) of 10 years, living in west Germany with a regional unemployment rate of 7%. We present the development of the the median employment rates together with bootstrapped 95% confidence intervals.

In Figure 3 we shown employment rates over the life cycle by the initial combination of employment and health status (i.e at the age of 29). Conditional on being employed in the

Table 2: Estimation results: men aged 30-59

	Specification 1		Specification 2	
	Coeff.	St. error	Coeff.	St. error
Poor health				
Years of education/10	-0.98	(0.08)	-1.43	(0.13)
Age/100	18.34	(2.90)	18.14	(3.96)
$\text{Age}^2/100$	-16.02	(3.17)	-13.53	(4.35)
Number of children	-0.10	(0.02)	-0.17	(0.03)
Initial poor health	1.14	(0.05)	2.17	(0.08)
Initial share of experience	-0.68	(0.21)	-1.22	(0.31)
Lagged poor health (t-1)	2.23	(0.04)	1.19	(0.06)
Lagged non-employment (t-1)	0.44	(0.06)	0.21	(0.08)
Constant	-5.93	(0.63)	-6.06	(0.87)
Mass point 1			2.20	(0.10)
Mass point 2			-0.51	(0.12)
Non-employment				
Years of education/10	-1.71	(0.10)	-2.12	(0.12)
Age/100	-15.81	(3.22)	-20.61	(3.85)
$\text{Age}^2/100$	26.39	(3.55)	34.20	(4.26)
Number of children	-0.05	(0.02)	-0.09	(0.03)
Single	0.25	(0.05)	0.27	(0.07)
German nationality	-0.29	(0.07)	-0.41	(0.09)
Regional unemployment rate	4.96	(0.38)	6.08	(0.48)
Initial non-employment	0.81	(0.06)	1.54	(0.08)
Initial share of experience	-2.64	(0.23)	-3.31	(0.28)
Lagged non-employment (t-1)	3.24	(0.04)	2.65	(0.05)
Lagged poor health (t-1)	0.86	(0.06)	0.71	(0.07)
Constant	1.51	(0.69)	4.16	(0.85)
Mass point 1			-0.99	(0.11)
Mass point 2			-2.37	(0.11)
Unobserved heterogeneity				
π_1			0.23	(0.14)
π_2			1.19	(0.14)
Variance Poor health			1.19	(0.07)
Variance Non-employment			0.92	(0.09)
Correlation			0.52	(0.04)
Log likelihood	-14036.5		-13595	

Number of observations 26,999. Median number of time periods 6.11 (min =2, max= 11). The regional unemployment rate ("Arbeitslose nach Kreisen") varies by 438 areas. Initial share of experience is defined as the quotient of the reported working years and the age at the initial observation. π_1 and π_2 are transformed probabilities that describe the population average percentage at the discrete mass points. Standard errors of the variances and the correlation are derived using parametric bootstrap with 1000 replications.

Source: Authors' calculations based on GSOEP 1996-2007 data and Bundesanstallt für Arbeit.

Table 3: Estimation results: men aged 30-54

	Specification 1		Specification 2	
	Coeff.	St. error	Coeff.	St. error
Poor health				
Years of education/10	-1.02	(0.09)	-1.60	(0.14)
Age/100	15.85	(4.42)	14.81	(5.69)
Age^2/100	-12.70	(5.15)	-8.92	(6.65)
Number of children	-0.10	(0.02)	-0.15	(0.03)
Initial poor health	1.16	(0.05)	2.00	(0.09)
Initial share of experience	-0.67	(0.24)	-1.60	(0.34)
Lagged poor health (t-1)	2.23	(0.04)	1.28	(0.07)
Lagged non-employment (t-1)	0.54	(0.07)	0.38	(0.09)
Constant	-5.46	(0.92)	-4.88	(1.19)
Mass point 1			1.97	(0.11)
Mass point 2			-0.76	(0.13)
Non-employment				
Years of education/10	-1.77	(0.11)	-2.05	(0.15)
Age/100	-12.33	(5.06)	-5.10	(5.94)
Age^2/100	22.20	(6.01)	13.75	(7.05)
Number of children	-0.02	(0.03)	-0.10	(0.03)
Single	0.34	(0.06)	0.43	(0.08)
German nationality	-0.35	(0.08)	-0.37	(0.10)
Regional unemployment rate	5.29	(0.42)	5.61	(0.56)
Initial non-employment	0.80	(0.06)	1.77	(0.10)
Initial share of experience	-2.76	(0.26)	-2.48	(0.33)
Lagged non-employment (t-1)	3.12	(0.05)	2.51	(0.07)
Lagged poor health (t-1)	0.89	(0.07)	0.57	(0.09)
Constant	0.92	(1.04)	0.96	(1.23)
Mass point 1			-1.00	(0.12)
Mass point 2			-2.65	(0.12)
Unobserved heterogeneity				
π_1			-0.02	(0.16)
π_2			1.05	(0.14)
Variance Poor health			1.13	(0.08)
Variance Non-employment			1.10	(0.09)
Correlation			0.58	(0.04)
Log likelihood	- 11267		-10909	

Number of observations 23,175. Median number of time periods 6.12 (min =2, max= 11). For further notes, see Table 2.

Source: Authors' calculations based on GSOEP 1996-2007 data and Bundesanstallt für Arbeit.

initial state, employment rates for this stylised household a relatively high. Between the age of 30 and 50 the initial health status (conditional on being employed at 29) only marginally affects the employment risk. However in the last ten years the employment risk increases for those with initial poor health. At the age of 55 the difference in the median employment rate by health status amounts to 7 percentage points and increases to nearly 20 percentage points at the age of 59. This divergence is driven by the estimated persistence in the health and the endogeneity of the two risks.

The difference by initial health risk is similar when assuming that the man was not employed in the initial period. However, in this scenario the development of the employment rate over the life-cycle is very different independent of the initial health status. Due to the state dependence in employment, the probability of employment remains very low in the first years, reaches levels close to 90% around the age of 40 and significantly decreases towards the end of the working life. In particular the low levels after the age of 50 can be attributed to the individual effects captured by the initial employment state. Again in the beginning employment rates hardly differ by the initial health risk. However the risk of non-employment is always greater among those with initial poor health and it is nearly twice as large for those aged 59.

In Figure 4 we slightly change the simulation and allow the onset of a permanent health shock (i.e. a poor health risk of 100%) to vary by age. We simulate three scenarios, with onset of the permanent health shock at the age of 30, 40 or 50 and compare these to employment rates for the same stylised individual who never experiences poor health. For all simulations we assume that the man is employed in the initial state. The simulations show that employment rates significantly decrease when the individual is "hit" by the health shock. Again we find the largest difference after the age of about 55. Interestingly the magnitude of this difference is only minimally affected by the timing of the onset of poor health which is the effect of rapid convergence of the employment path when conditioned on permanent poor health.

4 Conclusion

The relationship between health and labour market activity is a very important one from the point of view of socio-economic policy. Understanding how much poor health limits employment prospects and to what extent labour market activity or inactivity affects health can help in determining the role of potential health improvements on employment levels and indicate the degree to which the latter can contribute to the changes in health status.

In this paper we have developed a joint intertemporal model of poor health and non-employment risks in which we allow for correlation of unobservables and mutual dependence of the two processes. The model has been estimated on a sample of German men aged 30-59 using the SOEP data for the years 1996-2007.

The results confirm a significant interaction between the two outcomes and demonstrate persistence in both processes. Poor health at *(t-1)* is shown to be a significant determinant of current non-employment risk and lagged non-employment has a positive effect on the probability of being in poor health at time (t). Our results highlight the importance of controlling for unobserved heterogeneity in the estimation. Accounting for unobserved effects reduces the magnitude of the estimated coefficients on the lagged endogenous variables and significantly reduces the persistence of both processes. Simulations based on our results demonstrate the importance of the initial health and employment conditions for the development of the employment rate over the life cycle and show that the onset of permanent poor health significantly lowers employment probability. Health plays an especially strong role in determining labour market outcomes among those aged over 50. In our stylised example permanent poor health reduces the employment rate at the age of 59 by around 35 percentage points and this effect only marginally depends on the timing of the onset of permanent poor health.

Our findings stress the importance of health for labour market outcomes throughout the life-cycle and its particular role among those aged over 50. The effects of health on employment are significant even after controlling for the correlated unobserved heterogeneity which affects both of the examined processes. Employment in later stages of life also seems to be very strongly affected by individual characteristics as captured by the initial health and employment states. Our analysis confirms earlier findings on the role of employment in the early years of the working life on the later development of the employment path, and shows that health status in these early years will also play an important role in determining employment dynamics. The findings thus point out towards a potentially crucial role of health-related policies throughout working lives and underline the importance of employment support for younger individuals. Both types of policies can lead to substantial gains in labour market activity particularly in the later stages of individuals' lives.

References

Adams, P., M. D. Hurd, D. McFadden, A. Merrill, and T. Ribeiro (2003): "Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status," *Journal of Econometrics*, 112, 1, 3–56.

Akay, A. (2009): "The Wooldridge Method for the Initial Values Problem Is Simple: What About Performance," *IZA Discussion Paper*, 3943.

Alessie, R., S. Hochguertel, and A. van Soest (2004): "Ownership of stocks and mutual funds: a panel data analysis," *Review of Economics and Statistics*, 86, 783–796.

Benitez-Silva, H., and H. Ni (2008): "Health Status and health dynamics in an empirical model of expected longevity," *Journal of Health Economics*, 27, 564–584.

Blundell, R., C. Meghir, and S. Smith (2002): "Pension incentives and the pattern of early retirement," *Economic Journal*, 112, C153–C170.

Bratberg, E. (1999): "Disability Retirement in a Welfare State," *The Scandinavian Journal of Economics*, 101(1), 97–114.

Carrasco, R. (2001): "Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the Effect of Fertility on Female Labor Participation," *Journal of Business & Economic Statistics*, 19(4), 385–94.

Chan, S., and A. H. Stevens (2001): "Job Loss and Employment Patterns of Older Workers," *Journal of Labor Economics*, 19(2), 484–521.

Clark, A., and A. Oswald (1994): "Unhappiness and Unemployment," *Economic Journal*, 104, 648–659.

Gruber, J. (2000): "Disability Insurance Benefits and Labor Supply," *Journal of Political Economy*, 108(6), 1162–1183.

Heckman, J., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, No. 2, 271–320.

Hyslop, D. (1999): "State Dependence, serial correlation and hetergeity in intertempral labor force participation of married women," *Econometrica*, 67, 1255–1294.

Kalwij, A., and F. Vermeulen (2008): "Health and Labour Force Participation of Older People in Europe: What Do Objective Health Indicators Add to the Analysis?," *Health Economics*, 17(5), 619–638.

Kuhn, P., and A. Sweetman (1999): "Vulnerable Seniors: Unions, Tenure and Wages Following Permanent Job Loss," *Journal of Labor Economics*, 17(4), 671–693.

Manton, Kenneth, G. (1998): "Economic Effects of Reducing Disability," *American Economic Review*, 88(2, Papers and Proceedings), 101–105.

Michaud, P.-C., and K. Tatsiramos (2008): "Fertility and Female Employment Dynamics in Europe: The Effect of Using Alternative Econometric Modeling Assumptions," *IZA Discussion Paper*, 3853.

Riphahn, R., T. (1999): "Disability Retirement among German Men in 1980s," *Industrial and Labor Relations Review*, 52(4), 628–647.

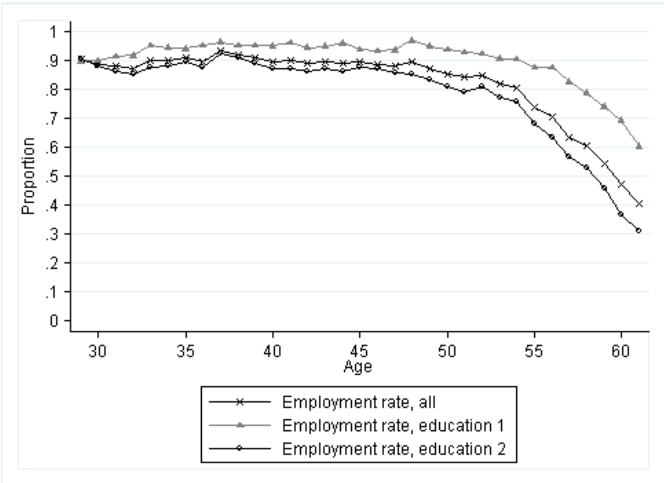
Shelton Brown III, H., J. Pagan, and B. E. (2005): "The impact of diabetes on employment: genetic IVs in bivariate probit," *Health Economics*, 14, 537–544.

Wooldridge, J. M. (2005): "Simple solution to the initial conditions problem in dynamic, nonlinear panel-data models with unobserved heterogeneity," *Journal of Applied Econometrics*, 20, 39–54.

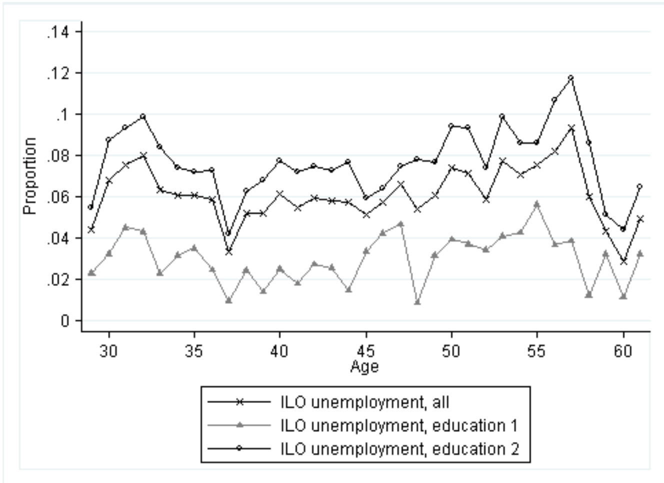
Figures

Figure 1: Labour market and health outcomes by age

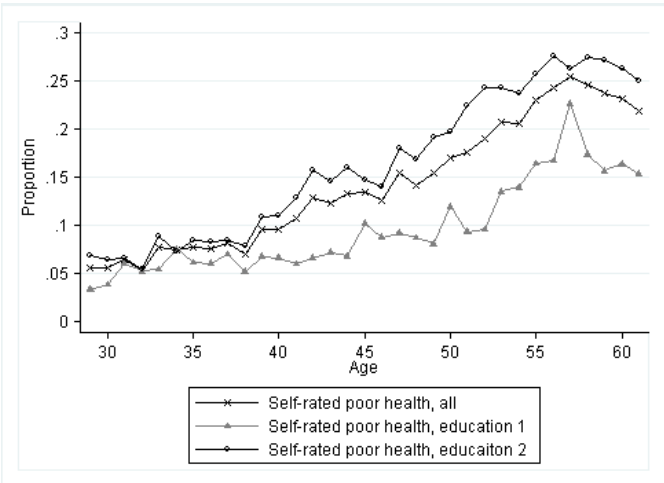
1A - Employment rate by age



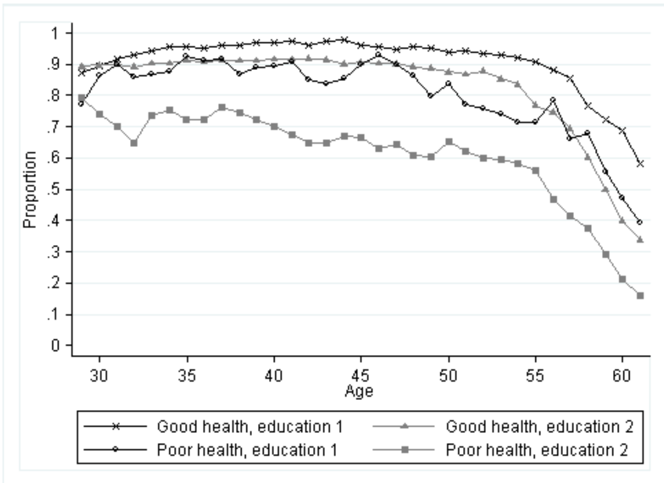
1B - ILO unemployment by age



1C - Self-rated poor health by age



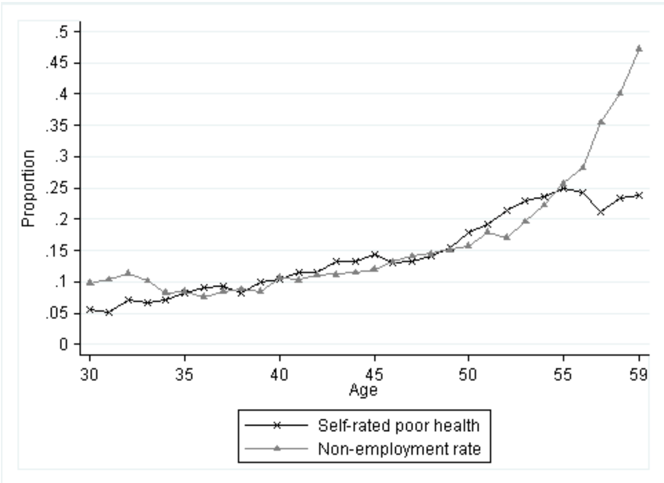
1D - Employment by health status



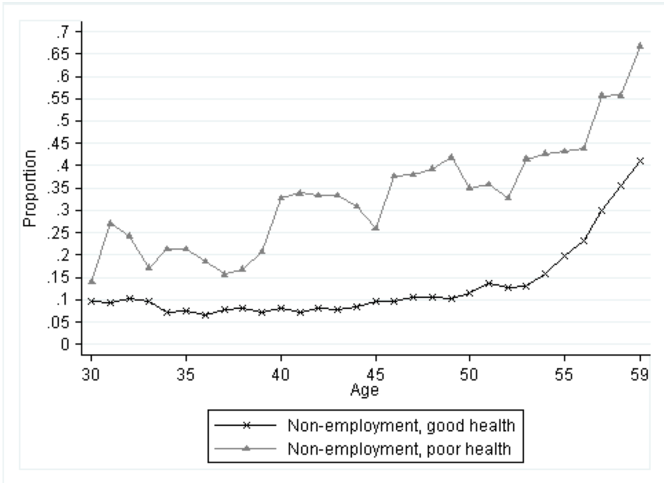
Source: Authors' calculations based on GSOEP 1996-2007 data. Full sample.

Figure 2: Estimation sample - poor health and non-employment

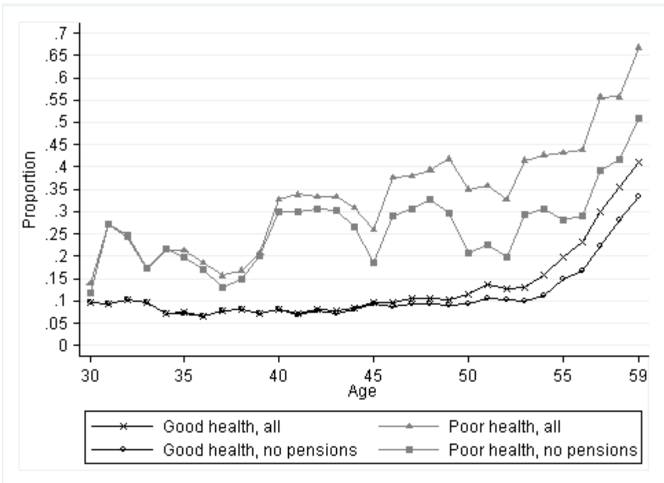
2A - Poor health and non-employment



2B - Non-employment conditional on health



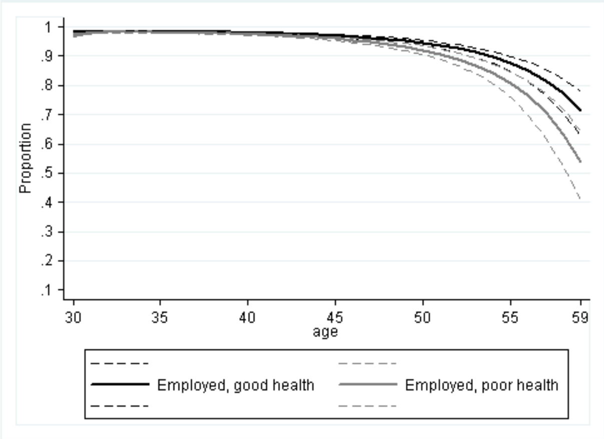
2C - Non-employment conditional on health and pension status



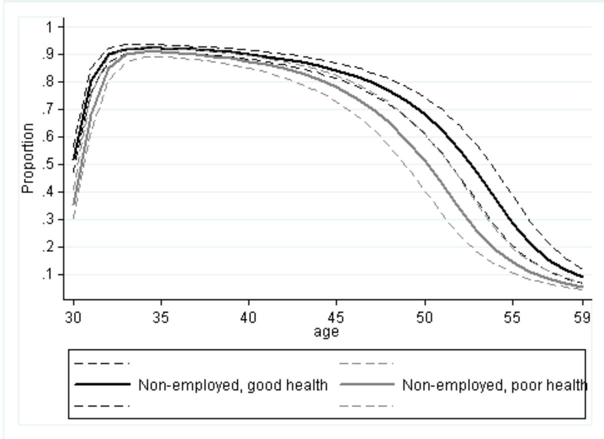
Source: Authors' calculations based on GSOEP 1996-2007 data. Estimation sample.

Figure 3: Employment simulations conditional on employment and health at age 29

3A - Employment path conditional on being employed at age 29



3B - Employment path conditional on being non-employed at age 29

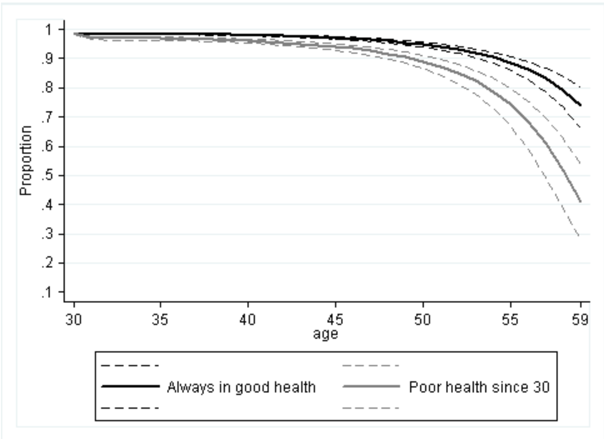


Source: Authors' calculations.

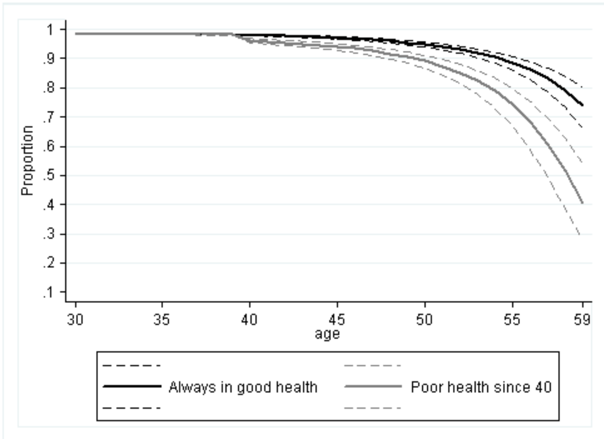
Notes: Dashed lines represent the respective 95% confidence intervals which are derived by using a parametric bootstrap with 1000 replications.

Figure 4: Employment simulations conditional on the timing of onset of poor health

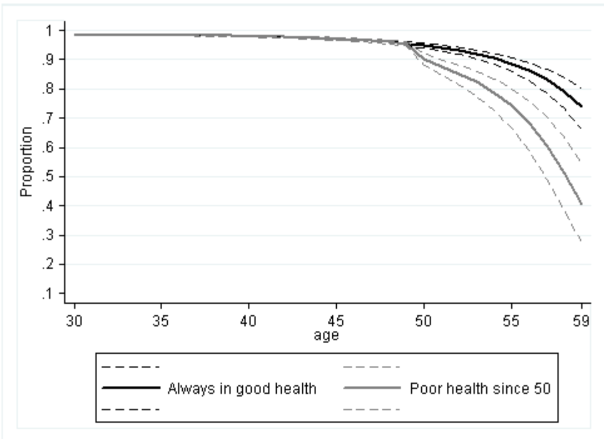
4A - Onset of (permanent) poor health at age of 30



4B - Onset of (permanent) poor health at age of 40



4C - Onset of (permanent) poor health at age of 50



Source: Authors' calculations.

Notes: Dashed lines represent the respective 95% confidence intervals which are derived by using a parametric bootstrap with 1000 replications.