Netspar

Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

The relationship between health and working full-time and part-time in old age

Evidence from a panel data simultaneous equations model

Paul Emsbroek

MSc 08/2018-008

The relationship between health and working full-time and part-time in old age: evidence from a panel data simultaneous equations model

Paul Emsbroek [443447]
B.Sc., University of Groningen 2017

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Econometrics & Mathematical Economics

Supervisor
Dr. Tunga Kantarcı

August 3, 2018

Abstract

Endogeneity is a major concern when modelling the relationship between health and labor market decisions. Health might be endogenous to labor due to justification bias or due to simultaneity. We employ a simultaneous equations model to analyze the causal effects of health and working in old age, allowing for endogeneity. In addition, we exploit the time dimension of the panel data to control for unobserved heterogeneity. Unlike the previous studies, we distinguish between working full-time and part-time to analyze the potentially heterogenous effects of working different numbers of hours. The results confirm the finding in the literature that better self-reported health has a positive effect on the probability of working. As for the reverse effect, we find that working full-time has a negative effect while working part-time has a positive effect on the general health of the elderly.

1 Introduction

Population aging is expected to become one of the most significant social transformations of the twenty-first century, with implications for nearly all sectors of society, including labor markets[1]. As a direct result of the population aging, the share of older workers in the labor force is increasing. In response, aging societies have reformed their pension systems to promote later retirement (French and Jones, 2017). Such reforms, however, are only effective if the health of older workers permits them to continue working. Hence, the relationship between health and retirement decisions is of great concern to policy makers.

Numerous studies have investigated the relationship between health and retirement. What makes studying this relationship particularly complicated is that there is empirical evidence for a causal relationship in both directions. While there are many studies that find that retirement has an effect on health, the effect of health on retirement is also well established in the literature.

Whether retirement is beneficial or harmful for physical or mental health has been the subject of a long-standing debate. While some studies argue that retirement is a stressful event that can lead to emotional or mental stress during the retirement years (Eisdorfer and Wilkie, 1977; Macbride, 1976), others suggest that retirement eleviates work stress and might preserve overall health (Ekerdt et al., 1983; Kerkhofs et al., 1999). The majority of the studies that analyze the effect retirement on health outcomes do not distinguish between full-time and part-time work. The following papers are among the most preeminent in this particular strand of the literature. Charles (2004), Coe and Zamarro (2011), Bonsang et al. (2012), Mazzonna and Peracchi (2012), Mazzonna and Peracchi (2016) and Rohwedder and Willis (2010) consider self-rated overall health as well as objective health measures, such as the physical functioning limitations, cognitive abilities, and depression, in the United States and in Europe. Bound and Waidmann (2007), Johnston and Lee (2009) and Behncke (2012) consider physical health measures, chronic conditions, and respondent self-reports of health in the United Kingdom. Eibich (2015) considers self-reported physical and mental health in Germany. Insler (2014) considers overall health in Europe and the United States and compares the results based on this health measure to those based on the health index proposed by Bound (1991). Hallberg et al. (2015) consider inpatient care and mortality in Sweden, and Biró (2016) considers outpatient care in the United States and Europe. Although there is some heterogeneity in the findings of these studies, this literature suggests that retirement is beneficial to health.

Studies have paid much less attention to the effects of working full-time and part-time on health. Vaus et al. (2007) find that Australian respondents making a gradual transition into retirement rate their overall health as being more positive compared to those who make an abrupt transition into retirement. Forbes et al. (2015) find that part-time workers rate their mental health as better than retirees rate in Australia. Dave et al. (2008) find that the physical and mental health of individuals who are partially retired are better than those of their retired peers in the United States. Neuman (2008) investigates the effects of reducing work hours to less than full-time in the United states, and finds that retiring or reducing work hours to less than full-time preserves physical health. As in the literature analyzing the effect of retirement on health, the main difficulty in identifying the effects of full-time and part-time work on health is that the labor market participation decisions at the extensive and intensive margins are potentially endogenous to health. Whereas Vaus et al. (2007) and Forbes et al. (2015) do not control for this potential bias, Dave et al. (2008) tackle the endogeneity problem by restricting their sample only to those individuals who had no major health problems and did not report deteriorating health prior to (partial) retirement. However, this empirical strategy is prone to selection bias. Neuman (2008) uses instrumental variables for the labor participation decision,

[1] United Nations – World population prospects, 2017 revision

which are dummies for reaching the retirement eligibility ages. Social security regulations in the United States offer significant financial benefits when retiring at specific retirement ages. Hence, retirement eligibility is a strong predictor of labor-force status. However, an additional difficulty is that the effects of working full-time and part-time need to be identified independent of each other. This means that we need instruments that provide independent sources of exogenous variation for full-time work and for part-time work, a concern that is not addressed in the existing literature.

In many respects, health is regarded as an important determinant of the labor market participation decision. Health can be viewed as a form of human capital, partly determining an individual's wage in the labor market (Becker et al., 1964; Grossman, 1972). Health shocks can affect the preferences for work and leisure, for example through their impact on the life expectancy (Chirikos, 1993; Dwyer and Mitchell, 1999). In the early life-cycle models, health was already recognized as one of the key factors affecting the retirement decision (Diamond and Hausman, 1984; Bazzoli, 1985). Although most of the early literature seems to agree that health has a significant effect on the retirement decision (Anderson and Burkhauser, 1984, 1985; Sickles and Taubman, 1986), the size of the effect of health in comparison to the effects of economic variables has been subject to discussion (Anderson and Burkhauser, 1984; Bazzoli, 1985). Bound (1991) argues that the large variation in effect sizes is due to the sensitivity of the labor supply models to the measures of health used in these models. When subjective measures of health, such as the self-reported health status, are used, health plays a more important role compared to when objective measures of health, such as chronic illness or physical functioning limitations, are used (Dwyer and Mitchell, 1999).

A main difficulty in estimating the effect of health on retirement is that health may be endogenous to labor market behavior. One source of endogeneity is due to the simultaneity of the health and labor market behavior. This may happen indirectly through various mechanisms. For example, decisions such as the investment in health care may depend on income, which in terms may depend on labor market participation. The decision whether to engage in physical exercise can depend on time available for leisure which can also depend on the number of hours worked (Grossman, 1972). In addition, labor market participation may directly affect health. For example, lack of activity for non-participants or stress associated with work might deteriorate health (Stern, 1989; Sickles and Taubman, 1986).

Another source of endogeneity of health to labor market participation decisions is due to measurement error in health. In many survey datasets, the only available measures of health are self-reported. Although there is a large number of studies indicating that self-reported measures are a significant and independent predictor of mortality (Okun et al., 1984; Idler and Kasl, 1995; Lundberg and Manderbacka, 1996), subjective measures of health may reflect ex-post rationalizations of retirement (Bazzoli, 1985). For example, individuals out of the labor force might report poor health to rationalize their non-participation. This type of endogeneity is often referred to as the justification endogeneity (Anderson and Burkhauser, 1984, 1985; Stern, 1989; Bound, 1989; Dwyer and Mitchell, 1999). Empirical evidence on justification endogeneity is not conclusive, however. Whereas Parsons (1982), Anderson and Burkhauser (1984), Anderson and Burkhauser (1985) suggest that justification endogeneity might exist, Stern (1989), Cai and Kalb (2006) do not find supporting evidence.

Several studies have used objective measures of health, such as mortality (Parsons, 1982; Anderson and Burkhauser, 1984) or parental mortality (Haan and Myck, 2009) to account for the justification endogeneity. Another popular approach in the literature is using objective health measures as instruments for self-reported health. Stern (1989) uses symptoms and diseases as instruments. Bound (1991), Bound et al. (1999), Dwyer and Mitchell (1999), and Disney et al. (2006) use indices constructed from chronic diseases or functional limitations. Campolieti (2002)

uses health conditions and BMI as instruments. However, objective health measures themselves are not free of problems. Bound (1991) argues that such measures are not perfectly correlated with different aspects of health that affect labor market decisions, leading to measurement error.

In addition to the endogeneity resulting from simultaneity between health and labor force status, and the justification endogeneity, Cai (2010) identifies another source of endogeneity, resulting from unobserved heterogeneity. Cai argues that unobserved factors might affect both health and labor market participation, correlating the error terms of the health and labor participation equations. For example, unobserved individual preferences for leisure might have effect on the number of working hours and might also have an indirect effect on health.

The review of the literature on the endogeneity of the health and labor market behavior suggests that health and labor market behavior need to be modelled jointly to control for different sources of endogeneity and to obtain unbiased estimates of the impact of one on another. Few studies have considered models that account for the different sources of endogeneity. Stern (1989) and Cai and Kalb (2006) utilize the simultaneous equations model using cross-sectional data. Sickles and Taubman (1986) and Cai et al. (2007) use panel data, and model the relationship between health and labor force status, but only allow health to affect the labor participation decision, and assume that the effect of labor force status on health is zero. Cai (2010) uses a simultaneous equations model with panel data but only considers working-age Australian men and women and does not distinguish between full-time and part-time work. In this study we will build on the earlier work of Stern (1989), Cai et al. (2007) and Cai (2010) and jointly model the relationship between health labor force participation. The contribution of our study to the literature on the relationship between health and retirement is threefold. First, studies analyzing the relationship between retirement on health model the labor market behavior as a binary decision of retirement and working and ignore the role of working different numbers of hours. We distinguish between working full-time and part-time and analyze their relationships with overall health in old age.

Second, we jointly model the relationship between health and retirement by employing a simultaneous equations model. We control for the endogeneity resulting from simultaneity and allow for correlation between the health equation and the labor participation equations. This modelling approach allows us to estimate both the causal effect of health on retirement and the reverse effect of retirement on health. Moreover, it allows us to conduct a 'true' test on the exogeneity of full-time work, part-time work and health by testing the joint significance of the parameters pertaining to the different sources of endogeneity. Within the simultaneous equations model, we employ an ordered probit specification to model the health equation, where self-reported health is the dependent variable, similar to Cai (2010). We use objective health measures constructed from physical functioning limitations and chronic illnesses as instruments for the potentially endogenous self-reported health, an approach that is common in the literature (Stern, 1989; Bound et al., 1999; Disney et al., 2006). In addition, we instrument health by using information on subsequent mortality of the respondent. Subsequent mortality often appears in the literature as a proxy for health (Anderson and Burkhauser, 1984, 1985). For the labor equation, we employ a multinomial-probit model to distinguish between three alternative labor market outcomes: full-time work, part-time work[2] and full-time retirement. We consider a number of instruments for labor market status, such as the retirement eligibility ages, following the many studies in the literature analyzing the effect of retirement on health Charles (2004); Neuman (2008); Coe and Zamarro (2011); Bonsang et al. (2012); Eibich (2015); Mazzonna and Peracchi (2016).

Third, we exploit the variation in unobserved taste-shifters across panel respondents over

[2]We define part-time work as working at most 34 hours per week, or as working at least 35 hours per week and at most 35 weeks per year, including the hours worked at a possible second job.

time by allowing the error of the regression model to follow a random effects structure. This allows us to control for unobserved heterogeneity, resulting in more efficient estimation results compared to when cross-sectional data are used.

The results show that being in good health increases the probability of working in old age. On the other hand, working full-time in old age is found to have a negative effect on health whereas working part-time in old age is found to have a positive effect on health. Furthermore, we find that controlling for random effects greatly improves efficiency of the estimates compared to when error components are assumed to be independent over time. In addition, we find that allowing for correlation between the error components of the equations for full-time and part-time work improves efficiency.

This paper is organized as follows. Section 2 describes the econometric model and the estimation approach. Section 3 describes the data. Section 4 presents the estimation results and discusses the validity of the instruments using a formal hypothesis test. Section 5 discusses the sensitivity of the results to alternative model specifications. Section 6 concludes.

2 Methodology

2.1 The model

We model labor market participation decision as a utility-based choice between three possible (exclusive) states $j = 1, 2, 3$. Here, $j = 1$ indicates full-time work, $j = 2$, indicates part-time work and $j = 3$ indicates retirement. With regards to the choice between the three possible working states, we assume that, in each period, choices for the three labor market states are made according to the random utility maximization hypothesis (Borsch-Supan et al., 1992), that is

$$i \text{ is chosen in period } t \text{ if } U^l_{it} > U^l_{jt} \text{ for } i, j = 1, 2, 3 \text{ where } i \neq j.$$

The utility of alternative j in period t is given by,

$$U^l_{jt} = \gamma_j U^{h^*}_t + x_{lt}\phi_{lj} + u_{jt} \text{ for } j = 1, 2, 3. \tag{1}$$

U^l_{jt} is the utility of labor force status j in time period t, which depends on true health utility $U^{h^*}_t$ and a vector of variables denoted by x_{lt}. We expect u_{jt} to be correlated over time as well as over the alternatives. The former is because unobserved factors affecting labor force utility can persist over time, and the latter because unobserved factors might exist and affect multiple choice alternatives, correlating their respective error terms. We allow for correlation over time by assuming that the error terms follow a random-effects structure.[3] The errors are then distributed as follows,

$$u_{jt} = \alpha_j + \varepsilon_{jt}, \alpha_j \sim \mathcal{N}(0, \sigma^2_{\alpha_j}), \varepsilon_{jt} \sim \mathcal{N}(0, \sigma^2_{\varepsilon_j}) \text{ for } j = 1, 2, 3$$

where α_j is i.i.d. over individuals and ε_{jt} is i.i.d. over individuals and over time.

We denote the vector of error components for all alternatives in all time periods as $u = (u_{11}, ..., u_{1t}, u_{21}, ..., u_{2t}, u_{31}, ..., u_{3t})$ with covariance matrix Ω_u.

[3] Imposing a fixed effects specification in this nonlinear model can be difficult due to the incidental parameters problem (Heckman, 1981). Moreover, identification of the fixed effect probit model relies on within-individual variation in the dependent and explanatory variables. Variables that do not vary over time can hence not be included in the estimation. In our sample, within-individual variation for labor-force status is limited, and therefore, a fixed effects estimation is not likely to lead to efficient estimates

The covariances between the error terms that appear in this matrix are specified as

$$\text{cov}(u_{js}, u_{it}) = \begin{cases} \sigma^2_{\alpha_j} + \sigma^2_{\varepsilon_j} \text{ if } i = j \text{ and } s = t \\ \sigma^2_{\alpha_j} \text{ if } i = j \text{ and } s \neq t \\ \sigma_{\alpha_j,\alpha_i} + \sigma_{\varepsilon_j,\varepsilon_i} \text{ if } i \neq j \text{ and } s = t \\ \sigma_{\alpha_j,\alpha_i} \text{ if } i \neq j \text{ and } s \neq t \end{cases} \tag{2}$$

The structure of this covariance matrix allows the error terms of the alternatives to be correlated with one another.

Since we are only interested in relative utilities, we write the utilities in (1) in relative terms by considering utility differences with respect to some base alternative (Geweke et al., 1994). In particular, we subtract the utility of retirement from the utility of each alternative as

$$\begin{aligned} \tilde{U}^l_{jt} = U^l_{jt} - U^l_{3t} &= \gamma^h_j U^{h*}_t + x_{lt}\phi_{lj} + u_{jt} - \gamma^h_3 U^{h*}_t - x_{lt}\phi_{l3} - u_{3t} \\ &= (\gamma^h_j - \gamma^h_3) U^{h*}_t + x_{lt}(\phi_{lj} - \phi_{l3}) + u_{jt} - u_{3t} \\ &= \tilde{\gamma}^h_j U^{h*}_t + x_{lt}\tilde{\phi}_{lj} + \tilde{u}_{jt} \text{ for } j = 1, 2. \end{aligned} \tag{3}$$

The vector of error differences in all time periods is denoted by $\tilde{u} = (\tilde{u}_{11}, ..., \tilde{u}_{1t}, \tilde{u}_{21}, ..., \tilde{u}_{2t})$ with covariance matrix $\Omega_{\tilde{u}}$. The covariances between the error differences can be written in terms of the covariances of the original error components as

$$\text{cov}(\tilde{u}_{js}, \tilde{u}_{it}) = \begin{cases} \sigma^2_{\alpha_j} + \sigma^2_{\alpha_3} - 2\sigma_{\alpha_j,\alpha_3} + \sigma^2_{\varepsilon_j} \\ \qquad + \sigma^2_{\varepsilon_3} - 2\sigma_{\varepsilon_j,\varepsilon_3} \text{ if } i = j \text{ and } s = t \\ \sigma^2_{\alpha_j} + \sigma^2_{\alpha_3} - 2\sigma_{\alpha_j,\alpha_3} \text{ if } i = j \text{ and } s \neq t \\ \sigma_{\alpha_j,\alpha_i} + \sigma^2_{\alpha_3} - \sigma_{\alpha_j,\alpha_3} - \sigma_{\alpha_i,\alpha_3} \\ + \sigma_{\varepsilon_j,\varepsilon_i} + \sigma^2_{\varepsilon_3} - \sigma_{\varepsilon_j,\varepsilon_3} - \sigma_{\varepsilon_i,\varepsilon_3} \text{ if } i \neq j \text{ and } s = t \\ \sigma_{\alpha_j,\alpha_i} + \sigma^2_{\alpha_3} - \sigma_{\alpha_j,\alpha_3} - \sigma_{\alpha_i,\alpha_3} \text{ if } i \neq j \text{ and } s \neq t \end{cases} \tag{4}$$

The relationship between health utility and the labor force status utility differences is specified as

$$U^{h^*}_t = \delta_1 \tilde{U}^l_{1t} + \delta_2 \tilde{U}^l_{2t} + x_{ht}\beta_h + u_{h^*t} \tag{5}$$

$U^{h^*}_t$ is latent true health at time t, which depends on the labor force status utility differences at time t (endogeneity of true health) and on a vector of variables denoted by x_{ht}.

However, we do not observe true health and therefore we are interested in how it is related to self-reported health. According to the justification hypothesis, respondents out of the labor force report worse health to justify their non-participation (Anderson and Burkhauser, 1984). We use U^h_t to represent the underlying subjective health utility which drives the observed self-reported health.

$$U^h_t = U^{h^*}_t + \lambda_1 \tilde{U}^l_{1t} + \lambda_2 \tilde{U}^l_{2t} + u_{h^{**}t} \tag{6}$$

We can now obtain an expression for the subjective health utility in terms of the labor force status utilities by substituting (5) into (6).

$$\begin{aligned} U^h_t &= \delta_1 \tilde{U}^l_{1t} + \delta_2 \tilde{U}^l_{2t} + x_{ht}\beta_h + u_{h^*t} + \lambda_1 U^l_{1t} + \lambda_2 U^l_{2t} + u_{h^{**}t} \\ &= (\delta_1 + \lambda_1)\tilde{U}^l_{1t} + (\delta_2 + \lambda_2)\tilde{U}^l_{2t} + x_{ht}\beta_h + u_{h^*t} + u_{h^{**}t} \\ &= \theta^l_1 \tilde{U}^l_{1t} + \theta^l_2 \tilde{U}_{2t} + x_{ht}\beta_h + u_{ht} \end{aligned} \tag{7}$$

Where δ represents endogeneity of true health and λ represents justification endogeneity. Since no information on true health status is available in the data, only θ, the sum of both sources of endogeneity, can be observed. However, given that the justification hypothesis predicts that the sign for λ is positive, we can infer that a negative θ must mean that endogeneity of true health must be negative and larger than the justification endogeneity. In case θ is positive, it can be that either the endogeneity of true health is positive or that true health endogeneity is negative but dominated by the justification endogeneity (Cai, 2010).

We now rewrite (6) as $U_t^{h*} = U_t^h - \lambda_1 \tilde{U}_{1t}^l - \lambda_2 \tilde{U}_{2t}^l - u_{h^{**}t}$ and substitute it into (3) to obtain an expression for the labor force status utilities in terms of the subjective health. We obtain

$$\begin{aligned}\tilde{U}_{1t}^l &= \tilde{\gamma}_1\big(U_t^h - \lambda_1 \tilde{U}_{1t}^l - \lambda_2 \tilde{U}_{2t}^l - u_{h^{**}t}\big) + x_{lt}\tilde{\phi}_{l1} + \tilde{u}_{1t} \\ &= \frac{\tilde{\gamma}_1}{(1-\tilde{\gamma}_1\lambda_1)}\tilde{\gamma}_1\big(U_t^h - \lambda_2 \tilde{U}_{2t}^l - u_{h^{**}t}\big) + x_{lt}\tilde{\phi}_{l1} + \tilde{u}_{1t} \end{aligned} \tag{8}$$

and similarly for $\tilde{U}_{2t}^l$

$$\begin{aligned}\tilde{U}_{2t}^l &= \tilde{\gamma}_2\big(U_t^h - \lambda_1 \tilde{U}_{1t}^l - \lambda_2 \tilde{U}_{2t}^l - u_{h^{**}t}\big) + x_{lt}\tilde{\phi}_{l2} + \tilde{u}_{2t} \\ &= \frac{\tilde{\gamma}_2}{(1-\tilde{\gamma}_2\lambda_2)}\big(U_t^h - \lambda_1 \tilde{U}_{1t}^l - u_{h^{**}t}\big) + x_{lt}\tilde{\phi}_{l2} + \tilde{u}_{2t} \end{aligned} \tag{9}$$

Subsequently, we substitute (8) into (9) and vice versa to obtain the equations for the labor utility differences expressed only in terms of subjective health and the vector of variables affecting the labor utilitites.

$$\begin{aligned}\tilde{U}_{1t}^l &= \frac{\tilde{\gamma}_1}{(1-\tilde{\gamma}_1\lambda_1)}\big(U_t^h - \lambda_2\big(\frac{\tilde{\gamma}_2}{(1-\tilde{\gamma}_2\lambda_2)}\big(U_t^h - \lambda_1 \tilde{U}_{1t}^l - u_{h^{**}t}\big) + x_{lt}\tilde{\phi}_{l2} + \tilde{u}_{2t}\big) - u_{h^{**}t}\big) + x_{lt}\tilde{\phi}_{l1} + \tilde{u}_{1t} \\ &= \frac{1}{1-\frac{\tilde{\gamma}_1\tilde{\gamma}_2\lambda_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)(1-\tilde{\gamma}_2\lambda_2)}} * \\ &\left(\frac{\tilde{\gamma}_1}{(1-\tilde{\gamma}_1\lambda_1)}\left(U_t^h - \lambda_2\Big(\frac{\tilde{\gamma}_2}{(1-\tilde{\gamma}_2\lambda_2)}\big(U_t^h - u_{h^{**}t}\big) + x_{lt}\tilde{\phi}_{l2} + \tilde{u}_{2t}\Big) - u_{h^{**}t}\right) + x_{lt}\tilde{\phi}_{l1} + \tilde{u}_{1t}\right) \\ &= \frac{1}{1-\frac{\tilde{\gamma}_1\tilde{\gamma}_2\lambda_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)(1-\tilde{\gamma}_2\lambda_2)}}\left(\frac{\tilde{\gamma}_1}{(1-\tilde{\gamma}_1\lambda_1)}\left(U_t^h - \lambda_2\Big(\frac{\tilde{\gamma}_2}{(1-\tilde{\gamma}_2\lambda_2)}\big(U_t^h\big) + x_{lt}\tilde{\phi}_{l2}\Big)\right) + x_{lt}\tilde{\phi}_{l1}\right) \\ &= \frac{\frac{\tilde{\gamma}_1}{(1-\tilde{\gamma}_1\lambda_1)}\left(1-\frac{\lambda_2\tilde{\gamma}_2}{(1-\tilde{\gamma}_2\lambda_2)}\right)}{1-\frac{\tilde{\gamma}_1\tilde{\gamma}_2\lambda_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)(1-\tilde{\gamma}_2\lambda_2)}}U_t^h + \frac{\tilde{\phi}_{l1} - \frac{\tilde{\gamma}_1\lambda_2\tilde{\phi}_{l2}}{(1-\tilde{\gamma}_1\lambda_1)}}{1-\frac{\tilde{\gamma}_1\tilde{\gamma}_2\lambda_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)(1-\tilde{\gamma}_2\lambda_2)}}x_{lt} \\ &+ \frac{\frac{\tilde{\gamma}_1}{(1-\tilde{\gamma}_1\lambda_1)}\left(\frac{\lambda_2\tilde{\gamma}_2}{(1-\tilde{\gamma}_2\lambda_2)}-1\right)}{1-\frac{\tilde{\gamma}_1\tilde{\gamma}_2\lambda_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)(1-\tilde{\gamma}_2\lambda_2)}}u_{h^{**}t} + \frac{\frac{-\tilde{\gamma}_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)}}{1-\frac{\tilde{\gamma}_1\tilde{\gamma}_2\lambda_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)(1-\tilde{\gamma}_2\lambda_2)}}\tilde{u}_{2t} + \frac{1}{\frac{\tilde{\gamma}_1\tilde{\gamma}_2\lambda_1\lambda_2}{(1-\tilde{\gamma}_1\lambda_1)(1-\tilde{\gamma}_2\lambda_2)}}\tilde{u}_{1t} \\ &= \theta_1^h U_t^h + x_{lt}\beta_{l_1} + u_{l_1t} \end{aligned} \tag{10}$$

and in a similar fashion,

$$\tilde{U}_{2t}^l = \theta_2^h U_t^h + x_{lt}\beta_{l_2} + u_{l_2t} \tag{11}$$

As can be observed in (10), the error term u_{l_1t} is a linear combination of the error terms in (6) and (3), meaning that even if (3) and (5) are assumed to be independent, the error term of (7)

is still correlated with the error terms in (10) and (11) through justification endogeneity. This means that, even if we assume the equations in (3) to be independent of each other, the error terms of (10) and (11) are still correlated through justification endogeneity.

In (10) and (11), we do not observe $\tilde{\gamma}$ and λ. Instead, we only observe their sum, θ^h. We cannot estimate the underlying parameters that make up the error terms u_{ht}, $u_{l_1 t}$ and $u_{l_2 t}$. Assuming a random effects structure for the error component of the health equation, u_{ht}, the error terms jointly follow a multivariate normal distribution with the covariance of the error terms specified as follows.

$$\text{cov}(u_{l_j s}, u_{l_i t}) = \begin{cases} \omega^2_{\alpha_j} + \omega^2_{\varepsilon_j} \text{ if } i = j \text{ and } s = t \\ \omega^2_{\alpha_j} \text{ if } i = j \text{ and } s \neq t \\ \omega_{\alpha_j,\alpha_i} + \omega_{\varepsilon_j,\varepsilon_i} \text{ if } i \neq j \text{ and } s = t \\ \omega_{\alpha_j,\alpha_i} \text{ if } i \neq j \text{ and } s \neq t \end{cases} \tag{12}$$

$$\text{cov}(u_{l_j s}, u_{ht}) = \begin{cases} \omega_{\alpha_h,\alpha_j} + \omega_{\varepsilon_h,\varepsilon_j} \text{ if } s = t \\ \omega_{\alpha_h,\alpha_j} \text{ if } s \neq t \end{cases} \tag{13}$$

$$\text{cov}(u_{hs}, u_{ht}) = \begin{cases} \omega^2_{\alpha_h} + \omega^2_{\varepsilon_h} \text{ if } s = t \\ \omega^2_{\alpha_h} \text{ if } s \neq t \end{cases} \tag{14}$$

In the above setting, the parameters are not all identified (Train, 2009). We have already normalized the multinomial probit model for size by expressing the labor force utilities as differences with respect to the third alternative and we have normalized the ordered probit model of health for size by excluding the intercept from the model (McKelvey and Zavoina, 1975). However, we must also normalize the labor force and health utilities for scale in order to identify the model. We normalize $\omega_{\varepsilon_1} = 1$ and $\omega_{\varepsilon_h} = 1$. This amounts to dividing $\tilde{U}^l_{1t}$ and $\tilde{U}^l_{2t}$ by ω_{ε_1} and dividing $\tilde{U}^h_t$ by ω_{ε_h}. This normalization further complicates the direct interpretation of the estimated coefficients since the new coefficients reflect the impact of the observed variables relative to the standard deviation of the time variant unobserved factors. In this normalized covariance matrix, fewer parameters are identified, but this is not a restriction. This covariance matrix contains fewer parameters because we simply eliminate irrelevant aspects of the original covariance matrix, namely the scale and the level of utility. The remaining parameters contain information about the variances and covariances of errors independent of scale and level (Train, 2009).

As an example, the variance-covariance matrix of the error terms of the structural equations in the case of 4 time periods is given by

$$\begin{aligned} &\text{cov}(u_{l1,1}, u_{l1,2}, u_{l1,3}, u_{l1,4}; u_{l2,1}, u_{l2,2}, u_{l2,3}, u_{l2,4}; u_{h,1}, u_{h,2}, u_{h,3}, u_{h,4}) \\ &\equiv \Omega = \begin{pmatrix} I_4\omega^2_{\varepsilon_1} + e_4 e_4' \omega^2_{\alpha_1} & I_4\omega_{\varepsilon_1,\varepsilon_2} + e_4 e_4' \omega_{\alpha_1,\alpha_2} & I_4\omega_{\varepsilon_h,\varepsilon_1} + e_4 e_4' \omega_{\varepsilon_h,\varepsilon_1} \\ I_4\omega_{\varepsilon_1,\varepsilon_2} + e_4 e_4' \omega_{\alpha_1,\alpha_2} & I_4\omega^2_{\varepsilon_2} + e_4 e_4' \omega^2_{\alpha_2} & I_4\omega_{\varepsilon_h,\varepsilon_2} + e_4 e_4' \omega_{\varepsilon_h,\varepsilon_2} \\ I_4\omega_{\varepsilon_h,\varepsilon_1} + e_4 e_4' \omega_{\alpha_h,\alpha_1} & I_4\omega_{\varepsilon_h,\varepsilon_2} + e_4 e_4' \omega_{\alpha_h,\alpha_2} & I_4\omega^2_{\varepsilon_h} + e_4 e_4' \omega^2_{\alpha_h} \end{pmatrix} \end{aligned} \tag{15}$$

where I_4 denotes the 4 by 4 identity matrix and e_4 is a vector of ones.

The latent variables as specified above are linked to their observed discrete counterparts in the following way.

For self-reported health h_t we employ an ordered-probit model,

$$h_t = \begin{cases} 4 \text{ (=excellent)} & \text{if } m_3 < U_t^h < +\infty \\ 3 \text{ (=very good)} & \text{if } m_2 < U_t^h \leq m_3 \\ 2 \text{ (=good)} & \text{if } m_1 < U_t^h \leq m_2 \\ 1 \text{ (=fair)} & \text{if } m_0 < U_t^h \leq m_1 \\ 0 \text{ (=poor)} & \text{if } -\infty < U_t^h \leq m_0 \end{cases} \tag{16}$$

where the cut-off points $m_0 ... m_3$ are estimated from the data.

For labor force status l_t^i, along the lines of the multinomial probit model,

$$l_t^i = 1 \text{ (i is chosen in period t) if } \tilde{U}_{jit}^l < 0 \ \forall j \neq i \text{ for } i, j = 1, 2, 3 \tag{17}$$

where i indicates the chosen alternative and $\tilde{U}_{jit}^l$ is the difference between the utility of alternative j and the utilty of alternative i. Note that all $\tilde{U}_{jit}^l$ can be computed from $\tilde{U}_{1t}^l$ and $\tilde{U}_{2t}^l$ (for example $\tilde{U}_{12t}^l = \tilde{U}_{1t}^l - \tilde{U}_{2t}^l$). This means that there is no need to include extra equations for the utility differences with respect to the utility of a different base alternative.

2.2 Estimation

2.2.1 The reduced forms

We estimate the model using the full-information maximum likelihood (FIML) method. The FIML method allows the error terms of the health equation and the labor market participation equations to be correlated, providing estimates for all parameters in the variance-covariance matrix of the three error terms. This allows us to conduct a 'true' test for the exogeneity of full-time work, part-time work and health by taking into consideration all potential sources of endogeneity.

In order to estimate the parameters of the simultaneous equations model, we must first rewrite the structural Equations (7), (10) and (11) in their reduced forms. The reduced forms are obtained by substituting one equation into the other. First we substitute (10) and (11) into (7) to obtain the reduced form equation of subjective health utility. We obtain

$$\begin{aligned} U_t^h &= \theta_1^l\big(\theta_1^h U_t^h + x_{lt}\tilde{\beta}_{l1} + \tilde{u}_{l1t}\big) + \theta_2^l\big(\theta_2^h U_t^h + x_{lt}\tilde{\beta}_{l2} + \tilde{u}_{l2t}\big) + x_{ht}\tilde{\beta}_h + \tilde{u}_{ht} \\ &= \big(\theta_1^l\theta_1^h + \theta_2^l\theta_2^h\big)U_t^h + \theta_1^l\big(x_{lt}\tilde{\beta}_{l1} + \tilde{u}_{l1t}\big) + \theta_2^l\big(x_{lt}\tilde{\beta}_{l2} + \tilde{u}_{l2t}\big) + x_{ht}\tilde{\beta}_h + \tilde{u}_{ht} \\ &= \frac{1}{1 - (\theta_1^l\theta_1^h + \theta_2^l\theta_2^h)}\Big(\theta_1^l\big(x_{lt}\tilde{\beta}_{l1} + \tilde{u}_{l1t}\big) + \theta_2^l\big(x_{lt}\tilde{\beta}_{l2} + \tilde{u}_{l2t}\big) + x_{ht}\tilde{\beta}_h + \tilde{u}_{ht}\Big) \\ &= \frac{1}{1 - (\theta_1^l\theta_1^h + \theta_2^l\theta_2^h)}\Big(x_{lt}\big(\theta_1^l\tilde{\beta}_{l1} + \theta_2^l\tilde{\beta}_{l2}\big) + x_{ht}\tilde{\beta}_h\Big) + \frac{1}{1 - (\theta_1^l\theta_1^h + \theta_2^l\theta_2^h)}\Big(\theta_1^l\tilde{u}_{l1t} + \theta_2^l\tilde{u}_{l2t} + \tilde{u}_{ht}\Big) \\ &= \chi_t\Lambda_h + \upsilon_{h,t} \end{aligned} \tag{18}$$

Now we can substitute (18) into (10) and (11) to obtain the reduced form equations for the

labor force status utility differences. We obtain

$$
\begin{aligned}
\tilde{U}^l_{1t} &= \frac{\theta^h_1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(\theta^l_1\big(x_{lt}\tilde{\beta}_{l1}+\tilde{u}_{l1t}\big)+\theta^l_2\big(x_{lt}\tilde{\beta}_{l2}+\tilde{u}_{l2t}\big)+x_{ht}\tilde{\beta}_h+\tilde{u}_{ht}\Big)+x_{lt}\tilde{\beta}_{l1}+\tilde{u}_{l1t}\\
&= \frac{\theta^h_1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(\theta^l_2\big(x_{lt}\tilde{\beta}_{l2}+\tilde{u}_{l2t}\big)+x_{ht}\tilde{\beta}_h+\tilde{u}_{ht}\Big)\\
&+ \frac{\theta^h_1\theta^l_1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\big(x_{lt}\tilde{\beta}_{l1}+\tilde{u}_{l1t}\big)+\frac{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\big(x_{lt}\tilde{\beta}_{l1}+\tilde{u}_{l1t}\big)\\
&= \frac{\theta^h_1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(\theta^l_2\big(x_{lt}\tilde{\beta}_{l2}+\tilde{u}_{l2t}\big)+x_{ht}\tilde{\beta}_h+\tilde{u}_{ht}\Big)+\frac{1-\theta^l_2\theta^h_2}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\big(x_{lt}\tilde{\beta}_{l1}+\tilde{u}_{l1t}\big)\\
&= \frac{1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(\big(1-\theta^l_2\theta^h_2\big)\big(x_{lt}\tilde{\beta}_{l1}+\tilde{u}_{l1t}\big)+\theta^h_1\theta^l_2\big(x_{lt}\tilde{\beta}_{l2}+\tilde{u}_{l2t}\big)+\theta^h_1\big(x_{ht}\tilde{\beta}_h+\tilde{u}_{ht}\big)\Big)\\
&= \frac{1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(x_{lt}\Big(\big(1-\theta^l_2\theta^h_2\big)\tilde{\beta}_{l1}+\theta^h_1\theta^l_2\tilde{\beta}_{l2}\Big)+x_{ht}\theta^h_1\tilde{\beta}_h\Big)\\
&+ \frac{1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(\big(1-\theta^l_2\theta^h_2\big)\tilde{u}_{l1t}+\theta^h_1\theta^l_2\tilde{u}_{l2t}+\theta^h_1\tilde{u}_{ht}\Big)\\
&= \chi_t\Lambda_{l1}+\upsilon_{l1,t} \text{ and similarly}
\end{aligned}
\tag{19}
$$

$$
\begin{aligned}
\tilde{U}^l_{2t} &= \frac{1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(\theta^h_2\theta^l_1\big(x_{lt}\tilde{\beta}_{l1}+\tilde{u}_{l1t}\big)+\big(1-\theta^l_1\theta^h_1\big)\big(x_{lt}\tilde{\beta}_{l2}+\tilde{u}_{l2t}\big)+\theta^h_2\big(x_{ht}\tilde{\beta}_h+\tilde{u}_{ht}\big)\Big)\\
&= \frac{1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(x_{lt}\Big(\theta^h_2\theta^l_1\tilde{\beta}_{l1}+\big(1-\theta^l_1\theta^h_1\big)\tilde{\beta}_{l2}\Big)+x_{ht}\theta^h_2\tilde{\beta}_h\Big)\\
&+ \frac{1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\Big(\theta^h_2\theta^l_1\tilde{u}_{l1t}+\big(1-\theta^l_1\theta^h_1\big)\tilde{u}_{l2t}+\theta^h_2\tilde{u}_{ht}\Big)\\
&= \chi_t\Lambda_{l2}+\upsilon_{l2,t}
\end{aligned}
\tag{20}
$$

The variance-covariance matrix of the reduced-form errors in equations (18), (19) and (20) in a 4 time period example can be written as

$$
\mathrm{cov}(\upsilon_{l1,1},\upsilon_{l1,2},\upsilon_{l1,3},\upsilon_{l1,4};\upsilon_{l2,1},\upsilon_{l2,2},\upsilon_{l2,3},\upsilon_{l2,4};\upsilon_{h,1},\upsilon_{h,2},\upsilon_{h,3},\upsilon_{h,4})\equiv\Omega^*=A\Omega A',
$$

$$
\text{where } A=\frac{1}{1-(\theta^l_1\theta^h_1+\theta^l_2\theta^h_2)}\begin{pmatrix} I_4(1-\theta^l_2\theta^h_2) & I_4\theta^h_1\theta^l_2 & I_4\theta^h_1\\ I_4\theta^h_2\theta^l_1 & I_4(1-\theta^l_1\theta^h_1) & I_4\theta^h_2\\ I_4\theta^l_1 & I_4\theta^l_2 & I_4\end{pmatrix}. \tag{21}
$$

The structural error components jointly follow a multivariate normal distribution. Hence, the probability of observing a given sequence of health-labor force states can be calculated using the reduced-form error terms. For example, the probability of observing the sequence

$$
(l_{11}=1,l_{22}=1,l_{33}=1,l_{14}=1;h_1=1,h_2=2,h_3=3,h_4=4)
$$

is

$$
\begin{aligned}
&P(l_{11}=1, l_{22}=1, l_{33}=1, l_{14}=1; h_1=1, h_2=2, h_3=3, h_4=4) \\
&= \int_{-\infty}^{-\chi_1\Lambda_{l21}} \int_{-\infty}^{-\chi_1\Lambda_{l31}} \int_{-\infty}^{-\chi_2\Lambda_{l12}} \int_{-\infty}^{-\chi_2\Lambda_{l32}} \int_{-\infty}^{-\chi_3\Lambda_{l13}} \int_{-\infty}^{-\chi_4\Lambda_{l23}} \int_{-\infty}^{-\chi_4\Lambda_{l21}} \int_{-\infty}^{-\chi_4\Lambda_{l31}} \\
&\int_{m_0-\chi_1\Lambda_h}^{m_1-\chi_1\Lambda_h} \int_{m_1-\chi_2\Lambda_h}^{m_2-\chi_2\Lambda_h} \int_{m_2-\chi_3\Lambda_h}^{m_3-\chi_3\Lambda_h} \int_{m_3-\chi_4\Lambda_h}^{m_4-\chi_4\Lambda_h} \\
&\phi\big(\upsilon_{l21,1}, \upsilon_{l31,1}, \upsilon_{l12,2}, \upsilon_{l32,2}; \upsilon_{l13,3}, \upsilon_{l23,3}, \upsilon_{l21,4}, \upsilon_{l31,4}; \upsilon_{h,1}, \upsilon_{h,2}, \upsilon_{h,3}, \upsilon_{h,4} | \Omega^*\big) \\
&\mathrm{d}\upsilon_{h,4}\mathrm{d}\upsilon_{h,3}\mathrm{d}\upsilon_{h,2}\mathrm{d}\upsilon_{h,1}\mathrm{d}\upsilon_{l21,1}\mathrm{d}\upsilon_{l31,1}\mathrm{d}\upsilon_{l12,2}\mathrm{d}\upsilon_{l32,2}\mathrm{d}\upsilon_{l13,3}\mathrm{d}\upsilon_{l23,3}\mathrm{d}\upsilon_{l21,4}\mathrm{d}\upsilon_{l31,4}
\end{aligned}
\tag{22}
$$

where $\tilde{U}^l_{jit} = \chi_t\Lambda_{jit} + \upsilon_{lji,t}$ is the difference between the utility of alternative j and the utility of chosen alternative i as in Equation (17). To estimate this probability using the conventional Maximum Likelihood estimation, an evaluation of all 12-dimensional integrals is required, making estimation infeasible. However, simulation-based estimation techniques provide a solution by replacing the integrals with random approximations. The details of this approach are discussed in the next section.

2.2.2 Maximum simulated likelihood

The probit probabilities laid out in Equation (22) must be approximated numerically as they do not have a closed-form expression (Train, 2009). There are several procedures that can be used in this situation. Quadrature methods can be helpful to approximate the the integral (Butler and Moffitt, 1982). However, such methods are not effective here because of the high dimension of the integral. An alternative is the Clark algorithm (Daganzo et al., 1977). As it turns out, however, accuracy of this approximation can be very low in certain situations (Horowitz et al., 1982). For the probit probabilities under consideration, simulation-based estimation has proved to be a very useful approach (Train, 2009). The likelihood function can be simulated in different ways. A common approach in the literature is the direct simulation of the likelihood function, known as maximum simulated likelihood. Following the notation of Hyslop (1999), let $\{\xi_i\} = \{\xi_{i1}, ..., \xi_{iR}\}$ be a sequence of primitive simulators, drawn independently of the model parameters and the data. Then,

$$
\tilde{L}(\theta; x_i, \xi_i) = \frac{1}{R}\sum_{r=1}^{R} \tilde{L}(\theta; x_i, \xi_{ir}) \tag{23}
$$

approximates the likelihood function for the unknown parameter vector θ given the random sample of observations $x_i, (i = 1, ..., N)$, if $\tilde{L}(\theta; x_i, \xi_{ir})$ is an unbiased simulator for this likelihood function. The maximum simulated likelihood estimator for θ is then defined as

$$
\hat{\theta}_{\mathrm{MSL}} = \arg\max_\theta \sum_{i=1}^{N} \ln\big(\tilde{L}(\theta; x_i, \xi_i)\big). \tag{24}
$$

The simulator that we employ in Equation (23) is the GHK simulator (Geweke et al., 1994; Hajivassiliou et al., 1996; Keane, 1994). Of the numerous simulators available, the GHK simulator is the most widely used probit simulator. Hajivassiliou et al. (1996) have found the GHK simulator to be the most accurate among many simulators in a wide variety of settings. The GHK simulator is both unbiased and consistent in the number of draws, bounded by zero and one and continuous in the parameters (Hajivassiliou, 1992). Although the simulated log-likelihood function in Equation (24) is biased for a finite number of replications due to the non-linear

logarithmic transformation, maximum simulated likelihood is consistent if $R \to \infty$ and $N \to \infty$ and is asymptotically efficient if $R/\sqrt{N} \to \infty$ (Hajivassiliou and Ruud, 1993).

In order to reduce variance of the simulation results, we use the method of antithetic acceleration (Geweke, 1988). Monte-Carlo evidence presented by Hajivassiliou (2000) suggests that multivariate normal rectangle probabilities are more accurately approximated using an antithetically accelerated simulator than using the standard GHK simulator. In order to obtain the maximum simulated likelihood estimator described in Equation (24) we use 250 uniform random draws together with their reflections, obtaining a total of 500 replications.[4]

3 Data

The data used in this study come from the Health and Retirement Study (HRS), a longitudinal panel study that surveys more than 20,000 people representative of the population of the United States. The survey is launched in 1992 and collects biennial information on various topics, including demographics, health, pensions and the financial situation at the household level.

We impose a number of restriction on the sample data. First, we drop respondents who either reported to never have worked, who reported to have worked but with a maximum tenure of 5 years on all jobs or for whom this information was missing in any of the survey years. Second, we drop the respondents who do not report having a job after the age of fifty in any of the survey years. Third, we drop those respondents who after reporting retirement in a survey year, reported to be either working, unemployed, disabled or not in the labor force in a subsequent survey year. This assures that retirement is an absorbing state. Fourth, all respondents are dropped who reported to be unemployed, disabled or not in the labor force in a given survey year since we analyze only labor market participation and retirement. Fifth, we drop respondents who were younger than 50 or older than 75 years old in a given survey year.

To the best of our knowledge, no procedures exist to employ the simultaneous equations model described in Section 2.1 for unbalanced panels. For this reason, we restrict the sample to individuals and subsequent waves where no observations are missing with respect to health and labor force status, for example, due to attrition. The four subsequent waves from 1998 to 2004 represent the largest balanced panel in terms of observations. The descriptive statistics of this balanced panel are reported in table 2.

3.1 Dependent variables

Self-reported health To collect information on the respondent's self-reported general health status, the respondent is asked the question, "Would you say your health is excellent, very good, good, fair, or poor?". We recode the response to this question so that higher values represent better health (0 = poor, 1 = fair, 2 = good, 3 = very good, 4 = excellent).

Full-time work In the HRS, the respondent is asked how many hours he or she usually works per week and how many weeks he or she usually works per year. We define full-time work as working 35 or more hours per week for 36 or more weeks per year, including hours worked at a second job.

[4] In various studies, it is found that using Halton draws instead of random draws can improve efficiency. However, the high-order integral under consideration requires primes of high order, for which correlations can exist. Although, this can be avoided by using scrambled Halton sequences, we do not find a significant improvement in efficiency.

Part-time work We define part-time work as working less than 35 hours a week, or as working more than 36 weeks a year, including hours worked at a second job.

Retired. We define retirement as working 0 hours per year, including hours worked at a second job.

3.2 Instrumental variables

The variables described in this section appear either in equation (7) or in equations (10) and (11). Excluded from equations (10) and (11) are the instrumental variables for health. Similarly, the labor instruments are excluded from equation (7). In this way, we satisfy the exclusion restrictions required to identify the simultaneous equations model and we do not have to rely exclusively on the non-linearity of the model for the parameters to be identified[5].

3.2.1 Instrumental variables for health

Health conditions In the HRS, respondents are asked whether a doctor has ever told the respondent that he or she has a particular (chronic) health condition. The eight included conditions are: high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychiatric problems, and arthritis. A dummy variable is created indicating whether the respondent has answered yes for any of the health conditions mentioned above. Although, strictly speaking, this variable is self-reported, we follow a suggestion by (Bound et al., 1995) that self-reported chronic health conditions can be treated as exogenous. Namely, more specific survey questions leave little room for subjective interpretation and are hence less prone to justification endogeneity.

Physical functioning In the HRS, respondents are asked if they have any difficulty performing one of the following tasks: bathing, eating, dressing, walking across a room, and getting in or out of bed. We construct a variable indicating the number of tasks that the respondent reports as difficult to perform.

Subsequent mortality The HRS records whether respondents pass away in any of the waves of the HRS dataset. Anderson and Burkhauser (1984) use subsequent mortality as an objective proxy for health status and argue that it is exogeneous to the labor-supply decision. Their argument is based on the fact that death is an absorbing state and can by itself never affect the retirement decision.

The use of the health conditions and physical functioning variables to instrument self-reported health is common in the literature. (Stern, 1989; Bound et al., 1995; Dwyer and Mitchell, 1999; Campolieti, 2002; Disney et al., 2006). Moreover, validity of such variables can be motivated by economic theory. For example, according to the human capital theory (Grossman, 1972), health conditions or limitations negatively affect health capital and, hence, only affect labor-force status indirectly through their effect on health capital.

3.2.2 Instrumental variables for labor-force status

Eligibility for social security benefits We consider three binary variables to indicate that the respondent has reached an age that makes him or her eligible for retirement benefits. In particular, the first variable indicates that the respondent has reached the age of 62, which is

[5] In contrast to classical linear models, multi-equation probit models with unrestricted error correlation structure are, in general, identified without any exclusion restrictions (Wilde, 2000)

the early retirement age when agents can claim their pension benefits. The second variable indicates that the respondent has reached the normal retirement age but is younger than 70 years old. The third variable is an indicator of being older than 70. The exact definition of the normal retirement age depends on the birth date of the respondent. The normal retirement ages for each birth cohort are displayed in Table 1. A large number of studies employ the retirement eligibility ages as indicators for labor-force status, showing that they are strong predictors of retirement. Moreover, in the analysis of the causal effect of retirement on health, these instruments are not likely to have a direct effect on health (Charles, 2004; Neuman, 2008; Rohwedder and Willis, 2010; Coe and Zamarro, 2011; Bonsang et al., 2012; Mazzonna and Peracchi, 2012, 2016). Inclusion of the indicator for having reached the age of 70 is based on two arguments. First, before the year 2000, Social Security in the United States withheld retirement benefits if the earnings of an individual eligible for normal retirement exceeded a certain level before having reached the age of 70. This policy, called the earnings test, may encourage retirees to return to work at the age of 70 or increase the number of working hours when reaching this age. Second, in the United States, social security benefits can be delayed until the age of 70. In return, individuals who retire at a later age are compensated in the form of increased retirement benefits at the age of retirement. This can motivate individuals to retire at the age of 70.

Eligibility for pension benefits claimed at the early and normal retirement ages create important incentives for many agents to retire from their full-time career job. In this respect, the retirement eligibility ages are strong predictors of the full-time work decision.

Furthermore, pension benefits can partially substitute the income from work, meaning that an individual does not need to work the same amount of hours to retain his past level of income. Based on this argument, Aaronson and French (2004) use the statutory retirement ages as instruments for part-time work. The retirement ages can affect the full-time and part-time decisions in various ways. Individuals are, for example, much more likely to retire at the normal retirement age than at the early retirement age, because pension benefits are penalized for early claming. On the other hand, as of the normal retirement age, social security regulations allow earning a limited income from work while drawing social security benefits. As a result, some individuals might choose to not retire fully but reduce their working hours upon reaching normal retirement age, in order to supplement their income. In addition, individuals who are already working part-time might choose to retire fully upon reaching the normal retirement age, substituting their part-time work income with full pension benefits.

Table 1: Retirement eligibility ages

Year of birth	Retirement eligibility ages		
	Early	Normal	Late
1937 or earlier	62	65	70
1938	62	65 and 2 months	70
1939	62	65 and 4 months	70
1940	62	65 and 6 months	70
1941	62	65 and 8 months	70
1942	62	65 and 10 months	70
1943-1954	62	66	70
1955	62	66 and 2 months	70
1956	62	66 and 4 months	70
1957	62	66 and 6 months	70
1958	62	66 and 8 months	70
1959	62	66 and 10 months	70
1960	62	67	70

Source: The United States Social Security Administration

Partner's eligibility for social security benefits We construct three variables indicating the retirement eligibility of the respondent's partner or spouse. The first variable indicates that the respondent's spouse has reached the age of 62, which is the early retirement eligibility age. The second variable indicates that the respondent's spouse has reached the normal retirement age but is younger than 70 years old. The third variable is an indicator of being older than 70. The instruments are not likely to have an effect on the respondent's health, but may explain the respondent's retirement decision. Henretta et al. (1993); Blau (1998), and Gustman and Steinmeier (2000a, 2004) show that couples often synchronize their retirement. Therefore, the retirement eligibility of the respondent's spouse seems a natural predictor of the retirement decision of the respondent.

There are many ways in which retirement of the spouse can affect one's own retirement decision. Individuals value retirement more once their partner is retired (Gustman and Steinmeier, 2000a). Furthermore, a decrease in household income due to the retirement of the partner may induce individuals to retire at a later age. Gustman and Steinmeier (2014) find that males are likely to work beyond the early retirement age if their partners are retired, but are likely to decrease their number of working hours.

Ever self-employed We define a dummy variable that indicates whether the respondent has ever been self-employed in any of the survey waves of the HRS dataset. We include this variable instead of a variable indicating current self-employment status, since the latter has missing observations for people who are retired and results in dropping respondents who retire at some point during the survey years. We consider self-employment as an instrumental variable for labor-supply because self-employed individuals might be more flexible in their working schedules than people who work for an employer (Ekerdt et al., 1996; Kim and DeVaney, 2005). Furthermore, there is no direct reason to expect that people who have ever been self-employed have a better or worse health outcomes than people have never been self-employed.

Ever able to reduce hours We construct a variable indicating that the respondent has

had the ability to reduce the number of working hours at some point in the HRS dataset. Individuals who are never able to reduce hours are those who cannot change from a full-time to a part-time job, making this variable a natural predictor for part-time work. It can be argued, however, that jobs offering the option to reduce hours are generally less demanding and that people in ill health are more likely to choose such less demanding jobs. On this ground, it is possible that this variable is endogenous to health.

Union membership In the HRS, respondents are asked whether they are covered by a union or employee-association contract. We define a variable that indicates whether the respondent has ever answered yes to this question. Ehrenberg et al. (1988) argue that part-time workers are less likely to vote for a union or be a union member than their full-time working peers. In this respect, union membership presents itself as a natural predictor of part-time work.

Past usual retirement age In all survey years of the HRS, respondents are asked "On your main job, what is the usual retirement age for people who work with you or have the same kind of job?". Brown (2006) finds significantly large spikes in retirement hazard around the reported usual retirement age, indicating that, around this age, retirement probability is considerably higher than around other ages. Some respondents report several usual retirement ages. In such cases, we use the average of the reported usual retirement ages. We construct a dummy variable indicating that the age of the respondent is larger than the reported usual retirement age.

Health insurance provided by employer The HRS contains several questions on the topic of health insurance. Respondents who have indicated having other insurance than Medicare, military health care plans or long-term care insurance are asked the following two questions, "Do you obtain this health insurance through your own business or professional organization/current employer?" and "Do you obtain this health insurance through a former employer of yours?" We construct a variable where we assign a value of 1 if the respondent answers 'yes' to at least one of these questions for at least one of his primary pensions.

French and Jones (2011) find that access to health insurance affects the retirement decision. They find that raising the Medicare eligibility age leads individuals to work longer. Rust and Phelan (1997) find that whether or not an individual has Medicare significantly affects the retirement decision. Based on these findings, we expect individuals with employer-provided health insurance to retire later in order to retain their insurance and we expect them to be mostly full-time workers, since full-time workers receive more benefits in general. A drawback of this variable is that individuals who have health insurance have better access to (preventive) health care. Therefore, it could be argued that having health insurance from the employer may be correlated with health.

3.3 Control variables

We control for a range of variables that are known to affect both the labor market participation decision and health of the respondent. Controlling for such factors is important if we want to avoid confounding and omitted variable bias, which can both arise when we do not include relevant variables that are correlated with the dependent variable and one or more covariates in the equation. Adding more relevant variables to the model also allows us to explain more of the observed variation in the outcome variable. Moreover, when performing a test for the validity of instrumental variables, controlling for external factors can be important as it may lead to stronger conclusions about the exogeneity of the instruments. We test this in Section 3.3. First, we control for the evolution of self-reported health and labor market decisions throughout the

biological age of the respondent. The main control variable being age. To control for the effects of age on health and labor-supply, we employ a quadratic functional form of age. It is a common finding in the literature that health deteriorates with age in a nonlinear fashion (Kenkel, 1995; Haan and Myck, 2009; Coe and Zamarro, 2011). The probability of labor market participation is also found to decrease nonlinearly as a function of age (Gustman and Steinmeier, 1984a,b; Honig and Hanoch, 1985).

Health differences between men and women are well documented in the literature. Among others, Neuman (2008); Cai (2010); Coe and Zamarro (2011) report that subjective health is higher among women than among men. Furthermore, Gustman and Steinmeier (2000b) and Cahill et al. (2015) report that the probability of full-time and part-time work differs significantly between men and women in the HRS dataset. Therefore, we control for the potential gender effects by including a gender dummy in our model.

Another control variable that we consider is education. A higher level of education may be beneficial to health due to greater awareness of the health risks based on better health-related knowledge and accordingly larger investments in health care. Additionally, Grossman (1999) shows that education is a key factor in health production. Gustman and Steinmeier (1984a); Honig and Hanoch (1985) and Ruhm (1990) report significant differences in full-time and part-time frequencies as a function of education.

Household size may affect the labor market participation decision. One might argue, for example, that a bigger household has higher costs, increasing the economic incentives to work more. On the other hand, having a large household might be an incentive to reduce working hours in order to spend more time with the children. Household size may also affect the health outcome. For example, Gove and Geerken (1977) find evidence that having more children negatively affects mental health.

A control variable often mentioned in the literature is whether the respondent performs manual labor requiring physical effort, often described as blue-collar work, or administrative labor, often referred to as white-collar work. Performing manual labor might deteriorate health through physical stress (Sickles and Taubman, 1986). Westerlund et al. (2009) show that in France, blue-collar workers, on average, retire earlier than white-collar workers.

Table 2: Descriptive statistics

Variable name	Mean value	Missing (%)
Endogeneous variables		
Self-reported health	2.48	
Working full-time	0.46	
Working part-time	0.18	
Retired	0.36	
Instrumental variables		
Has a physical functioning problem	0.10	
Has a health condition	0.74	
Dies in any of the subsequent waves	0.19	
Under age 62	0.46	
Between age 62 and normal ret. age	0.18	
Between normal ret. age and age 70	0.25	
Over age 70	0.10	
Ever observed self-employed	0.27	10.6
Ever able to reduce paid hours	0.63	27.4
Ever member of a union	0.27	11.1
Employer provided health-insurance	0.55	
Control variables		
Age in years	62.6	
Respondent is male	0.59	
Has at least a college degree	0.49	
Household size	2.47	
Blue-collar worker	0.37	1.2
N. waves	4	
N. total obs.	14076	

Note. Due to rounding, totals may not add up to 1.

4 Results

4.1 Validity of exclusion restrictions

In Section 3.2, we discuss why the health instruments are unlikely to affect the labor participation decision and why the labor instruments are unlikely to affect the self-reported health. Despite these theoretical arguments in favor of the imposed exclusion restrictions, there might be undiscussed mechanisms that cause the exclusion restrictions to be violated. The variables excluded in each labor equation or in the health equation are used to instrument the endogenous variable in that equation. Hence, testing whether the exclusion restrictions hold amounts to testing whether the excluded variables are valid instruments (Cai, 2010). Commonly used tests of overidentifying restrictions, such as the Sargan-Hansen test (Sargan, 1958; Hansen, 1982) are only suitable for linear models. These tests cannot be applied here because of the discrete nature of the dependent variables, resulting in a nonlinear model where each endogenous variable enters the equation as a latent variable. For simultaneous equations models with limited dependent variables, Lee (1992) introduces a test of overidentifying restrictions that

is obtained as a by-product of a minimum distance estimator proposed by Amemiya (1978). This test, however, is not clearly established for the multinomial probit model. Therefore, we take a different approach to test for the exogeneity of the instrumental variables. We test the exogeneity of an instrument for health by including the instrument as an explanatory variable in the labor equations which also include the endogenous health variables which are predicted by other instruments. If a health instrument is truly exogenous to the labor supply decision, the corresponding coefficient of the instrument in that equation must be zero. We use the Z-test to test whether the coefficient is zero. Failure to reject the null hypothesis indicates this. The results are reported in Table 3.

Table 3: Z-test scores for exogeneity of health instruments

	Health conditions		Physical functioning		Subsequent mortality	
	Full-time	Part-time	Full-time	Part-time	Full-time	Part-time
Z-statistic	−1.25	−1.40	−1.75	−1.18	−1.82	−1.66
$p \geq Z$	0.21	0.16	0.08	0.24	0.07	0.10

1. Standard errors used to calculate the Z-statistics are derived using the Newey-West sandwich estimator (Newey and West, 1986) and are robust to heteroskedasticity and autocorrelation.

The test provides some evidence against the exogeneity of the health instruments. This evidence is particularly strong for subsequent mortality, where the null hypothesis of exogeneity is rejected at the 10% significance level in both the full-time and the part-time equation. This can indicate that subsequent mortality is not exogenous to the market participation decisions, which is not in line with our expectation that subsequent mortality is most likely to be exogenous among the three health instruments. However, this exogeneity test hinges on the assumption that the two instruments included in the health equations are strong instruments of health. If this is not the case for health conditions and physical functioning, the test results for subsequent mortality might not be reliable and rejection of the exogeneity of subsequent mortality might indicate that, in fact, health conditions and physical functioning are poor instruments for health. For this reason, the results of this test are not conclusive.

To test the exogeneity of the set of instruments indicating retirement eligibility for the respondent and the set of instruments indicating retirement eligibility for the respondent's spouse, we use a test similar to the test for over-identified models outlined in Ettner (1996). She includes a set of potentially endogenous instruments into the second stage equation of an IV-regression and tests for joint significance of the instrument set. We use a similar approach, within the simultaneous equations framework.

We include the set of retirement eligibility age indicators for the respondent in the health equation, instrumenting the labor market participation decision using the retirement eligibility ages of the partner and test for the significance of the retirement eligibility ages of the respondent in the health equation. Then, we include the set of retirement eligibility ages of the partner in the health equation, instrumenting the labor market participation decisions using the retirement eligibility ages of the respondent and test for the significance of the set included in the health equation. If the set of instruments is not jointly significant in the health equation, we can maintain the assumption of exogeneity. We use the Wald test to test the null hypothesis that the three coefficients are all equal to zero. Failure to reject the null hypothesis indicates this. The results are reported in Table 4. Given the Wald statistics reported in the table, we have

no reason to reject the null hypothesis that the instrumental variables based on the retirement eligibility ages are exogenous.

Table 4: Wald-statistics for scores for exogeneity of labor instruments

	Health equation	
	Retirement eligibility ages	Retirement eligibility ages (P)
Wald statistic	1.51	1.80
$p \geq Wald$	0.68	0.61

1. Standard errors used to calculate the Z-statistics are derived using the Newey-West sandwich estimator (Newey and West, 1986) and are robust to heteroskedasticity and autocorrelation. 2. (P) stands for married or unmarried partner

4.2 Estimation results

Table 5 presents the estimation results of the baseline specification where we do not include any control variables other than a quadratic function of age. The full set of health instruments is used in the health equation, namely, an indicator for health conditions, an index variable that measures physical functioning of the respondent, and an indicator of subsequent mortality. We observe that all three health instruments are significant at the 1% level and the signs are negative, as would be expected. The age variable in our study is scaled for estimation purposes by subtracting 50 (the lowest age in the sample), and dividing by 10. The age variable included in the health equation is not significant at any of the conventional significance levels. The squared age variable, however, is significant at the 5% level and has a negative sign, suggesting that health is quadratic and concave in the age of the respondent. This is in line with the literature, where, in general, health is found to depreciate with age (Bound et al., 1999; Grossman, 1999).

In both labor equations, six instruments are considered. A set of three instruments indicate retirement eligibility of the respondent and a second set of three instruments indicate retirement eligibility of the respondent's spouse. In the full-time equation, early and normal retirement eligibility age indicators are significant at the 1% level and late retirement eligibility is significant at the 5% level. All three indicators have negative signs, implying that becoming eligible for retirement reduces the utility of full-time work compared to the utility of retirement. This finding suggests that full-time workers decrease their number of working hours when they become eligible for social security benefits, in line with many of the earlier studies (Charles, 2004; Neuman, 2008; Rohwedder and Willis, 2010; Coe and Zamarro, 2011; Bonsang et al., 2012; Mazzonna and Peracchi, 2012, 2016). The largest effect is found for the indicator of retirement eligibility at the normal age, which is plausible since pension benefits are not penalized at this age. The three instruments based on the retirement eligibility of the respondent's spouse are significant at the 1% level and all have a negative sign. The strongest effect is found for indicator of late retirement eligibility.

In the part-time equation, the indicators for early and normal retirement eligibility are significant at the 1% level and have a negative sign. This suggests that, when eligible for retirement, respondents favor retirement more than part-time work. The relative sizes of the instruments follow a similar pattern as in the full-time equation. Since in both equations utility is expressed as a difference from the reference retirement utility, we can directly compare the coefficients. The coefficients of the six eligibility age indicators are smaller in the part-time

equation than in the full-time equation. This means that upon reaching the retirement age, the utility for part-time work increases compared to the utility of full-time work, suggesting that, in general, retirement eligibility induces individuals to reduce their working hours, instead of working full-time. In both labor equations, the linear as well as the quadratic age variable are significant at the 1% level. The negative sign of the age variable and the positive sign of the squared age variable suggest that the effect of age on labor utility follows a downward convex quadratic function.

The effects of health on full-time and part-time work are positive and significant at the 1% level, and the effect sizes are similar. This implies that, ceteris paribus, individuals in good health have a larger probability of working than individuals in poor health. This result is in line with the common finding in the literature that health positively affects labor supply (Stern, 1989; Cai et al., 2007; Cai, 2010). The effect of health on the utility difference between full-time work and retirement is not significantly different from the effect of health on the utility difference between part-time work and retirement. This finding suggests that, using this specification, no evidence is found that health affects full-time and part-time work decisions differently.

In this model, the effects of the full-time and part-time work utilities on health are interpreted with respect to a reference utility which is for the retirement. Therefore, the negative coefficient for the full-time work utility (the difference between the full-time work utility and retirement utility) indicates that, ceteris paribus, an increase in the probability of full-time work relative to retirement decreases health. The positive coefficient of the part-time work utility (the difference between the part-time work utility and retirement utility) suggests that, ceteris paribus, an increase in the probability of part-time work relative to retirement increases health. The negative effect of full-time work utility is in line with Neuman (2008) who finds that retiring or reducing work hours to less than full-time preserves health. The positive effect of the part-time work utility is in line with Dave et al. (2008), who find that in the United States, individuals who are partially retired have better physical and mental health than retirees.

Due to the complicated nature of the model, it is not possible to calculate the marginal effects of health, working full-time, and working part-time. As an alternative, in Section 4.4 we use prediction to assess the effects of health, part-time work and full-time work.

The estimates of the variance and covariance parameters of the labor equations imply that unobserved heterogeneity plays an important role. The time-invariant error variance is relatively high compared to the variance of the time-variant error component. Moreover, the error terms of the full-time and part-time equations are highly and positively correlated. This is not surprising since the utility difference of full-time work and the utility difference of part-time work share one important commonality. They both represent the utility difference of working versus being retired. It is, therefore, not unlikely that many unobserved factors will affect both utility differences in the same direction. This finding supports the use of the multinomial probit model rather than, for example, a multinomial logit model which is much more restrictive in nature. By construction, the logit model imposes the Independence of Irrelevant Alternatives (IIA) assumption (Train, 2009), which is unlikely to hold here, given the large and significant estimate for the error covariance of the two labor equations. We discuss this in more detail in Section 5.2.

In addition to the direct effects of health on labor and vice versa, the two are also correlated indirectly through unobserved factors that affect both the health outcome and the labor market participation decision. This correlation is captured by the covariance parameters of the time-invariant errors, $\omega_{\alpha_h,\alpha_{FT}}$ and $\omega_{\alpha_h,\alpha_{PT}}$ and the covariance parameters of the time-varying errors, $\omega_{\varepsilon_h,\varepsilon_{FT}}$ and $\omega_{\varepsilon_h,\varepsilon_{PT}}$. The estimates for the covariance parameters of the time-invariant errors are not significant at any conventional significant levels. The covariance parameters of the time-varying errors are both negative and significant at the 1% level. The negative correlation

between the time-varying errors in the health and full-time work equations together with the negative effect of full-time work on health implies that treating health as exogenous could result in underestimation of the effect of health on full-time work.

Table 5: Results for model of health and full-time and part-time work

	Baseline equation	
	Coef	SE
Health equation		
Full-time work utility	-0.118^{*}	0.063
Part-time work utility	0.134^{*}	0.058
Age (minus 50)/10	-0.064	0.732
Age deviation squared/100	-0.174^{**}	0.085
Activities of daily living	-0.553^{***}	0.037
Health conditions	-0.978^{***}	0.064
Subsequent mortality	-0.759^{***}	0.07
Full-time equation		
Health	0.641^{***}	0.092
Constant	16.176^{***}	1.875
Age (minus 50)/10	-10.736^{***}	1.567
Age deviation squared/100	0.994^{***}	0.354
Between age 62 and normal ret. age	-1.199^{***}	0.225
Between normal ret. age and age 70	-1.621^{***}	0.333
Over age 70	-1.17^{**}	0.477
Between age 62 and normal ret. age (P)	-0.918^{***}	0.201
Between normal ret. age and age 70 (P)	-1.275^{***}	0.26
Over age 70 (P)	-2.102^{***}	0.382
Part-time equation		
Health	0.592^{***}	0.104
Constant	13.84^{***}	1.808
Age (minus 50)/10	-10.434^{***}	1.553
Age deviation squared/100	1.333^{***}	0.355
Between age 62 and normal ret. age	-0.797^{***}	0.219
Between normal ret. age and age 70	-0.938^{***}	0.321
Over age 70	-0.679	0.452
Between age 62 and normal ret. age (P)	-0.696^{***}	0.195
Between normal ret. age and age 70 (P)	-0.902^{***}	0.252
Over age 70 (P)	-1.273^{***}	0.349
Covariance parameters and cut-off estimates		
m_0	-4.797^{***}	0.79
m_1	-3.173^{***}	0.768
m_2	-1.497^{**}	0.744
m_3	0.31	0.718
$\omega_{\alpha_{FT}}$	6.841^{***}	0.811
$\omega_{\alpha_{PT}}$	6.504^{***}	0.8
$\omega_{\varepsilon_{PT}}$	0.897^{***}	0.104
ω_{α_h}	1.284^{***}	0.101
$\omega_{\alpha_{FT},\alpha_{PT}}$	42.958^{***}	0.564
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.43^{**}	0.187
$\omega_{\alpha_h,\alpha_{FT}}$	1.372	2.659
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	-0.532^{***}	0.123
$\omega_{\alpha_h,\alpha_{PT}}$	1.085	2.946
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	-0.608^{***}	0.122
Log-likelihood	23126.3	
No. of observations	3519	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. (P) stands for Married or unmarried partner. 3. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

4.3 Exogeneity of the health and labor-force status

Potential endogeneity does not only enter the model through the simultaneity and justification endogeneity measured by the parameters $\theta_1^h, \theta_2^h, \theta_1^l$ and θ_2^l, but also through the covariance parameters of the error terms in equations (7), (10) and (11). Using the full-information maximum likelihood estimates, a 'true' test on the exogeneity of the self-reported health and the labor-force status can be conducted, where we consider the coefficients corresponding to the different sources of endogeneity. If the full-time work decision were exogenous to health, the covariance of both the time-variant and time-invariant error components of the full-time work and health equations as well as the coefficient measuring the effect of health on the full-time, retirement utility difference (based on the full-time work utility and retirement utility difference) should all equal zero. By means of a Wald test, we test the joint significance of the set of three variables corresponding to the different sources of endogeneity of health in the full-time and the part-time work decisions. Equivalenty, we test the endogeneity of full-time work and part-time work in the health equation. The null hypothesis for exogeneity of, for example, full-time work is,

$$\theta_1^l = 0 \text{ , } \omega_{\alpha_h,\alpha_1} = 0 \text{ and } \omega_{\varepsilon_h,\varepsilon_1} = 0$$

Table 6 presents the test results. We observe that in all four cases, the Wald statistics are significant at any conventional significance level. This implies that self-reported health is endogenous to labor-force status and labor-force status is endogenous to self-reported health.

Table 6: Wald-statistics for test on exogeneity of health and labor

	Health equation		Full-time equation	Part-time equation
	Full-time	Part-time	Health	Health
Wald statistic	68.55	38.30	37.54	45.55

Note. The 1% critical value for the Chi-squared distribution with 3 degrees of freedom is $\chi^2_{3,0.99} = 11.34$.

4.4 Assessing the effects of health and labor

Due to the complicated nature of the model, marginal effects of health and labor cannot be computed. Therefore, we assess the effects of health on the work decisions by predicting the changes in the probabilities of working full-time and part-time conditional on the latent value of health. Our approach is similar to that of Cai (2010) in that we define five ranges of health by utilizing the cut-off estimates $m_0, ..., m_3$ corresponding to the five values of self-reported health status. We predict the probabilities of working full-time and working part-time conditional on these health values by averaging over all individuals in the sample.

Table 7 shows the predicted conditional probabilities of full-time and part-time work. The second column shows that, for example, an increase from fair to good health increases the probability of full-time work with about 3 percentage points. It appears that the effect of health on the probability of working full-time and part-time is non-linear. An increase from poor to fair health appears to result in a higher increase in the probability to work full-time than an increase from very good to excellent health. Furthermore, it appears that the effects of health on the probability of working part-time are smaller than those on the probability of working full-time.

Table 7: Predicted conditional probabilities of full-time and part-time work

	Full-time		Part-time	
Health status	Pred. Probability	Perc. point change	Pred. Probabilty	Perc. point change
Poor	0.4280		0.1575	
Fair	0.4599	3.1840	0.1662	0.8773
Good	0.4904	3.0493	0.1748	0.8528
Very good	0.5194	2.9000	0.1829	0.8167
Excellent	0.5468	2.7394	0.1906	0.7686

We assess the effects of the labor market participation decision on health by predicting the probabilities of reporting a specific health status conditional on combinations of the utility difference of full-time work and retirement and the utility difference of part-time work and retirement. According to equation (17), individuals choose to work full-time if the utility difference of full-time work and retirement is positive and larger than that of part-time work. They choose to work part time if the utility difference of part-time work and retirement is positive and larger than that of full-time work and retirement. They choose to retire if both utility differences are negative. We predict the probabilities of each self-reported health status conditional on ranges of the underlying utility differences that correspond to the three choice alternatives of labor market participation. Table 8 shows the predicted conditional probabilities for each of the five self-reported health outcomes and the predicted conditional health values. As expected, ceteris paribus, an individual who works full-time has a lower health than an individual who is retired. The opposite holds for part-time work.

In addition, we observe that the beneficial health effects of part-time work compared to full-time work and retirement are mainly due to having lower probabilities of poor and fair health and a substantially higher probability of having excellent health. The beneficial health effects of retirement compared to full-time work are more homogenous over the health outcomes.

Table 8: Predicted conditional probabilities of self-reported health

	Predicted probabilities					Predicted health
Labor force status	Poor	Fair	Good	Very good	Excellent	
Full-time work	0.1505	0.1761	0.2404	0.2374	0.1956	2.15
Part-time work	0.1167	0.1459	0.2171	0.2575	0.2628	2.40
Retirement	0.1418	0.1687	0.236	0.2437	0.2098	2.21

5 Sensitivity analysis

5.1 Instrumental variables and identification

In the literature examining the causal effect of retirement on health, indicators for reaching the retirement eligibility ages are commonly used instruments for labor market participation

decisions (Charles, 2004; Neuman, 2008; Rohwedder and Willis, 2010; Coe and Zamarro, 2011). In this study, we have extended this set of instruments with the retirement eligibility ages of the respondent's spouse. As shown in Table 5, the retirement eligibility ages of the respondent and the retirement eligibility ages of the partner are strong predictors of the full-time and part-time work decisions. Here, we investigate the sensitivity of the model estimates to restricting the instrument set to the retirement eligibility ages of the respondent, or to the retirement eligibility ages of the spouse. If our model is able to exploit the independent sources of exogenous variation offered by the two types of instruments and separately identify the effects of working full-time and part-time, we should observe that the effect of working full-time should be less precisely estimated if we restrict the instrument set to the retirement eligibility ages of the spouse, and the effect of working part-time should be less precisely estimated if we restrict the instrument set to the retirement eligibility ages of the respondent.

Table 9 presents the results. The main finding is that the effects of labor on health become insignificant when we restrict the instrument set to the retirement eligibility ages of the respondent and when we restrict the instrument set to the retirement eligibility ages of the spouse. This is probably a consequence of the reduced predictive power of the set. Moreover, the estimates of the retirement eligibility instruments of the respondent are less significant in the part-time equation than the retirement eligibility ages of the spouse, implying that the retirement eligibility ages of the spouse are better predictors of part-time work than the retirement eligibility ages of the respondent. We formally test between the restricted model and the unrestricted model by means of a log-likelihood test, where the unrestricted model includes all six instruments and each of the two restricted models includes only a set of three instruments. The likelihood-ratio test statistics are 72.2 and 85.6 respectively. We reject both the model using only the respondent's eligibility ages and the model using only the partner's eligibility ages, at any conventional significance level. These findings provide evidence that both sets of retirement ages provide independent sources of variation for working full-time and part-time, significantly improving the efficiency of the estimates.

A main concern for using the retirement eligibility ages as instruments is that they might be identifying the effect of working only for the subgroup of the population who face economic or non-economic incentives to retire at these ages. For example, self-employed individuals are not likely to be affected by the statutory retirement ages perhaps because they do not face mandatory retirement at the retirement ages from an employer. Therefore, the retirement eligibility instruments have a Local Average Treatment Effect interpretation (Imbens and Angrist, 1994). For this reason, the estimated effects might not reflect the labor market behavior of the subpopulations that base their labor supply decisions on factors other than the retirement age. We address this concern by employing a number of new instruments that might accommodate alternative labor market decision mechanisms. The new instruments we consider are the ability to reduce working hours, an indicator for having worked as self-employed, a variable indicating that the respondent is past the usual retirement age of his company, and an indicator of whether the respondent has health-insurance provided by his employer. These instruments are likely to capture the work preferences of sub-populations with broader or different incentives to work full-time or part-time instead of retirement.

We consider a number of different specifications for the instrument set. First, we replace the set of retirement eligibility ages of the respondent with the variable indicating that the respondent is past the usual retirement age in his company, or with the indicator of having health-insurance from the employer. Based on the arguments presented in Section 3.2, we expect these instruments to be strong predictors of the full-time work decision in particular. The results are presented in Table 10. Both newly introduced instruments are significant at the 1% level in the full-time and part-time equations and have the expected sign. The indicator

of having health-insurance from the employer has a substantially larger effect in the full-time equation than in the part-time equation, implying that this is a strong predictor of full-time work in particular. In both models, the effects of health on the utility differences of full-time and part-time work are positive and significant at the 1% level. The effects of full-time work and part-time work are not significant in these models. This could imply that either the findings reported in Table 5 are specific to the subgroups identified by the retirement eligibility age indicators or that the two newly introduced instruments for full-time work do not provide the exogenous sources of variation required to identify the effects of full-time and part-time work on health.

Second, we replace the set of retirement eligibility ages of the respondent's spouse with either the indicator of being able to reduce hours, or the indicator of ever being self-employed. Based on the arguments presented in Section 3.2, we expect that these instruments provide strong sources of exogenous variation for the part-time work decision. The results are presented in Table 11. In the model including the indicator of the ability to reduce working hours, the estimated effect of full-time work on health is negative and significant at the 10% level. The estimate of the effect of part-time work on health is positive and significant at the 5% level. The effects of health on full-time and part-time work are positive and significant at the 1% level. These findings are in line with those presented in Table 5. For the model including the indicator of ever being self-employed, the coefficient estimates of the effects of full-time and part-time work on health are insignificant. The effects of health on full-time and part-time work are positive and significant at the 1% level.

Third, we employ the new instruments to replace both the set of retirement eligibility ages of the respondent and the set of retirement eligibility ages of the spouse at the same time. We consider four combinations. In the first and second, we replace the retirement eligibility ages of the respondent by the indicator of ever being able to reduce working hours and replace the retirement eligibility ages of the partner either by the indicator of being past the usual retirement age, or by the variable indicating that the respondent receives health-insurance from the employer. In the third an fourth, we replace the retirement eligibility ages of the respondent by the indicator of ever being self employed and replace the retirement eligibility ages of the partner either by the indicator of being past the usual retirement age, or by the variable indicating that the respondent receives health-insurance from the employer. Tables 12 and 13 present the results. In all four combinations, the effects of health on full-time and part-time work are positive and significant at the 1% level, a finding that appears to be robust across all specifications considered in this section. The effects of full-time and part-time work are significant in the first and fourth specifications. In both specifications, the effect of full-time work on health is negative and the effect of part-time health is positive. These findings are in line with those presented in Table 5, implying that the negative effect of full-time and the positive effect of part-time work on health are not specific to the subpopulation of individuals retiring at the statutory retirement ages but also reflect the behavior of subpopulations that retire for different reasons.

The finding that health increases the probability of working full-time and part-time is robust across the different specifications discussed above. Across all specifications, we find that health has a positive effect on both full-time and part-time, significant at the 1% level. The significance of the estimates for the effect of full-time work and part-time work on health, however, varies over the specifications. Hence, we are interested in obtaining more conclusive evidence on the direction of the effects of full-time and part-time work on health. We argue that using the new instruments in addition to the sets of retirement eligibility ages of the respondent and the spouse might increase efficiency of the estimates due to the fact that different sources of exogenous variation are exploited. Furthermore, if addition of the instruments improves efficiency of

the estimates for the effect of full-time and part-time work on health, this implies that these effects are homogenous across the different subgroups in the population that retire for different reasons. We estimate the model using three new specifications. In each of these specifications, we supplement the set of six instruments pertaining to the retirement eligibility ages with one additional instrument. We consider the indicator for being able to reduce working hours and the variable indicating whether the respondent has ever reported to be self-employed. Additionally, we consider a variable indicating that the respondent has ever reported being member of a union, based on the argument of Ehrenberg et al. (1988) that part-time workers are less likely to vote for a union or be a union member than their full-time working peers. Table 14 presents the results. We find that introducing new instruments in addition to the retirement eligibility age indicators of the respondent and the partner improves the overall significance of the coefficient estimates. In all three specifications, the estimates for the retirement eligibility age indicators are significant and have the expected sign. Moreover, the effect of full-time work on health is found to be negative and the effect of part-time work is found to be positive, in line with what we find in Table 5. In the specification where we extend the labor instruments with the ability to reduce hours, the effects of full-time and part-time work on health are significant at the 5% level. In the specification where the indicator of ever being self-employed is used, the effects of full-time work and part-time work on health are significant at the 1% level. These findings support those of the baseline equation presented in Table 5 and provide some evidence that the results are not specific to the subgroup of individuals that retire due to incentives provided by the statutory retirement ages but are somewhat homogenous across different subgroups of the population retiring for different reasons.

Table 9: Restricted sets of retirement eligibility age indicators

	Ret. eligibility ages		Ret. eligibility ages (P)	
	Coef	SE	Coef	SE
Health equation				
Full-time work utility	−0.029	0.206	−0.054	0.225
Part-time work utility	0.017	0.31	0.046	0.305
Age (minus 50)/10	0.211	1.216	0.284	1.12
Age deviation squared/100	−0.208	0.241	−0.237	0.283
Activities of daily living	−0.544***	0.028	−0.566***	0.025
Health conditions	−1.016***	0.047	−1.014***	0.041
Subsequent mortality	−0.748***	0.062	−0.726***	0.06
Full-time equation				
Health	0.584***	0.104	0.506***	0.106
Constant	16.336***	1.915	17.462***	2.179
Age (minus 50)/10	−11.322***	1.572	−13.281***	1.847
Age deviation squared/100	0.875**	0.356	1.562***	0.353
Between age 62 and normal ret. age (res/par)	−1.299***	0.23	−1.107***	0.211
Between normal ret. age and age 70 (res/par)	−1.613***	0.33	−1.379***	0.271
Over age 70 (res/par)	−1.262***	0.481	−1.898***	0.371
Part-time equation				
Health	0.444***	0.121	0.364***	0.123
Constant	14.088***	1.849	15.338***	2.189
Age (minus 50)/10	−11.234***	1.575	−13.108***	1.857
Age deviation squared/100	1.334***	0.349	2.023***	0.346
Between age 62 and normal ret. age (res/par)	−0.905***	0.228	−0.846***	0.211
Between normal ret. age and age 70 (res/par)	−0.994***	0.326	−1.032***	0.271
Over age 70 (res/par)	−0.784*	0.467	−1.26***	0.355
Covariance parameters and cut-off estimates				
m_0	−4.528***	1.128	−4.492***	0.937
m_1.	−2.884**	1.123	−2.842***	0.933
m_2	−1.17	1.117	−1.15	0.928
m_3	0.653	1.112	0.675	0.923
$\omega_{\alpha_{FT}}$	6.933***	0.81	6.804***	0.854
$\omega_{\alpha_{PT}}$	6.858***	0.82	6.674***	0.869
$\omega_{\varepsilon_{PT}}$	0.822***	0.096	0.73***	0.095
ω_{α_h}	1.295***	0.209	1.291***	0.146
$\omega_{\alpha_{FT},\alpha_{PT}}$	46.044***	11.095	44.217***	11.57
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.419**	0.173	0.394**	0.17
$\omega_{\alpha_h,\alpha_{FT}}$	1.419	4.653	1.485	3.42
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	−0.547***	0.14	−0.492***	0.153
$\omega_{\alpha_h,\alpha_{PT}}$	1.666	5.339	1.658	3.927
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	−0.432***	0.167	−0.366**	0.142
Log-likelihood	23162.4		23169.1	
No. of observations	3519		3519	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. (res) stand for respondent. 3. (par) stands for married or unmarried partner. 4. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table 10: Retirement eligibility ages of the spouse and alternative full-time instrument

	Past usual ret. age		Empl. health insurance	
	Coef	SE	Coef	SE
Health equation				
Full-time work utility	−0.101	1.118	−0.107	0.11
Part-time work utility	0.117	1.168	0.14	0.163
Age (minus 50)/10	−0.039	0.63	−0.148	0.613
Age deviation squared/100	−0.105	0.181	−0.069	0.114
Activities of daily living	−0.647***	0.034	−0.564***	0.029
Health conditions	−0.958***	0.044	−0.986***	0.05
Subsequent mortality	−0.64***	0.072	−0.785***	0.064
Full-time equation				
Health	0.519***	0.095	0.675***	0.073
Constant	16.838***	1.54	13.856***	1.533
Age (minus 50)/10	−11.413***	1.358	−10.929***	1.351
Age deviation squared/100	1.793***	0.335	1.441***	0.272
Between age 62 and normal ret. age (P)	−0.796***	0.187	−0.828***	0.167
Between normal ret. age and age 70 (P)	−1.001***	0.248	−1.157***	0.222
Over age 70 (P)	−1.363***	0.355	−1.598***	0.297
Past usual ret. age / Empl. health insurance	−1.761***	0.236	1.262***	0.162
Part-time equation				
Health	0.504***	0.096	0.632***	0.08
Constant	16.455***	1.588	12.923***	1.572
Age (minus 50)/10	−11.386***	1.36	−10.595***	1.357
Age deviation squared/100	1.854***	0.326	1.63***	0.261
Between age 62 and normal ret. age (P)	−0.757***	0.19	−0.675***	0.169
Between normal ret. age and age 70 (P)	−0.958***	0.253	−0.992***	0.221
Over age 70 (P)	−1.249***	0.364	−1.234***	0.29
Past usual ret. age / Empl. health insurance	−1.686***	0.241	0.71***	0.165
Covariance parameters and cut-off estimates				
m_0	−4.669***	0.537	−4.669***	0.668
m_1.	−3.007***	0.544	−3.03***	0.67
m_2	−1.283**	0.552	−1.341**	0.672
m_3	0.52	0.559	0.477	0.674
$\omega_{\alpha_{FT}}$	5.764***	0.56	5.844***	0.65
$\omega_{\alpha_{PT}}$	5.761***	0.562	5.608***	0.657
$\omega_{\varepsilon_{PT}}$	0.983***	0.026	0.779***	0.08
ω_{α_h}	1.234***	0.088	1.313***	0.103
$\omega_{\alpha_{FT},\alpha_{PT}}$	33.17***	6.463	32.435***	7.479
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.972***	0.039	0.676***	0.13
$\omega_{\alpha_h,\alpha_{FT}}$	0.895	1.827	1.071	1.797
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	−0.526***	0.098	−0.704***	0.073
$\omega_{\alpha_h,\alpha_{PT}}$	0.937	1.903	1.187	1.804
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	−0.504***	0.107	−0.638***	0.081
Log-likelihood	17207.7		22986.8	
No. of observations	2584		3519	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. (P) stands for Married or unmarried partner. 3. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table 11: Retirement eligibility ages of respondent and alternative part-time instrument

	Able to reduce hours		Ever self-employed	
	Coef	SE	Coef	SE
Health equation				
Full-time work utility	-0.127^{*}	0.074	-0.237	0.379
Part-time work utility	0.134^{**}	0.067	0.247	0.375
Age (minus 50)/10	0.166	0.233	0.147	0.173
Age deviation squared/100	-0.217^{***}	0.052	-0.132^{**}	0.053
Activities of daily living	-0.609^{***}	0.032	-0.576^{***}	0.028
Health conditions	-0.929^{***}	0.044	-0.917^{***}	0.04
Subsequent mortality	-0.664^{***}	0.069	-0.71^{***}	0.062
Full-time equation				
Health	0.47^{***}	0.157	0.455^{***}	0.164
Constant	16.961^{***}	2.237	16.12^{***}	0.938
Age (minus 50)/10	-13.74^{***}	2.14	-11.52^{***}	1.027
Age deviation squared/100	1.259^{***}	0.453	0.872^{**}	0.382
Between age 62 and normal ret. age	-1.528^{***}	0.307	-1.459^{***}	0.243
Between normal ret. age and age 70	-1.82^{***}	0.425	-1.657^{***}	0.354
Over age 70	-1.689^{***}	0.636	-1.544^{***}	0.549
Able to reduce hours / ever self-employed	5.554^{***}	0.86	5.472^{***}	0.45
Part-time equation				
Health	0.409^{**}	0.161	0.394^{***}	0.164
Constant	13.743^{***}	2.178	15.619^{***}	0.964
Age (minus 50)/10	-13.447^{***}	2.118	-11.529^{***}	1.027
Age deviation squared/100	1.746^{***}	0.448	0.999^{***}	0.375
Between age 62 and normal ret. age	-1.095^{***}	0.31	-1.386^{***}	0.242
Between normal ret. age and age 70	-1.138^{***}	0.429	-1.555^{***}	0.355
Over age 70	-1.176^{*}	0.631	-1.493^{***}	0.549
Able to reduce hours / ever self-employed	6.562^{***}	0.938	5.596^{***}	0.459
Covariance parameters and cut-off estimates				
m_0	-4.575^{***}	0.358	-4.383^{***}	0.289
m_1.	-2.938^{***}	0.363	-2.71^{***}	0.292
m_2	-1.229^{***}	0.373	-1.014^{***}	0.297
m_3	0.575	0.384	0.783^{**}	0.304
$\omega_{\alpha_{FT}}$	7.172^{***}	1.011	6.993^{***}	0.545
$\omega_{\alpha_{PT}}$	6.704^{***}	0.989	6.924^{***}	0.539
$\omega_{\varepsilon_{PT}}$	1.065^{***}	0.149	1.003^{***}	0.031
ω_{α_h}	1.224^{***}	0.035	1.223^{***}	0.025
$\omega_{\alpha_{FT},\alpha_{PT}}$	46.618^{***}	13.685	48.351^{***}	7.517
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.589^{***}	0.222	0.983^{***}	0.039
$\omega_{\alpha_h,\alpha_{FT}}$	0.88	0.938	1.041	0.697
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	-0.435^{***}	0.166	-0.444^{***}	0.167
$\omega_{\alpha_h,\alpha_{PT}}$	0.518	0.759	1.038	0.669
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	-0.49^{***}	0.165	-0.456^{***}	0.167
Log-likelihood	16902.7		20913.7	
No. of observations	2554		3146	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table 12: Alternative full-time instruments and able to reduce hours

	Past usual ret. age		Empl. health insurance	
	Coef	SE	Coef	SE
Health equation				
Full-time work utility	-0.191^{**}	0.094	-0.064	0.069
Part-time work utility	0.195^{**}	0.088	0.07	0.069
Age (minus 50)/10	-0.082	0.223	-0.03	0.149
Age deviation squared/100	-0.119^{**}	0.047	-0.051	0.04
Activities of daily living	-0.619^{***}	0.03	-0.587^{***}	0.026
Health conditions	-0.975^{***}	0.041	-0.956^{***}	0.037
Subsequent mortality	-0.694^{***}	0.064	-0.703^{***}	0.058
Full-time equation				
Health	0.578^{***}	0.09	0.487^{***}	0.09
Constant	16.713^{***}	1.499	13.837^{***}	1.107
Age (minus 50)/10	-13.873^{***}	1.447	-13.602^{***}	1.238
Age deviation squared/100	1.868^{***}	0.313	1.711^{***}	0.262
Able to reduce hours	4.622^{***}	0.515	5.182^{***}	0.491
Past usual ret. age / Empl. health insurance	-1.926^{***}	0.235	1.761^{***}	0.183
Part-time equation				
Health	0.516^{***}	0.102	0.457^{***}	0.094
Constant	13.784^{***}	1.57	12.922^{***}	1.206
Age (minus 50)/10	-13.273^{***}	1.438	-13.266^{***}	1.257
Age deviation squared/100	2.148^{***}	0.305	1.794^{***}	0.256
Able to reduce hours	5.464^{***}	0.583	5.496^{***}	0.5
Past usual ret. age / Empl. health insurance	-1.602^{***}	0.233	1.333^{***}	0.203
Covariance parameters and cut-off estimates				
m_0	-4.772^{***}	0.333	-4.411^{***}	0.179
m_1.	-3.131^{***}	0.343	-2.747^{***}	0.176
m_2	-1.449^{***}	0.355	-1.064^{***}	0.175
m_3	0.338	0.367	0.734^{***}	0.175
$\omega_{\alpha_{FT}}$	5.996^{***}	0.609	5.75^{***}	0.516
$\omega_{\alpha_{PT}}$	5.675^{***}	0.599	5.694^{***}	0.519
$\omega_{\varepsilon_{PT}}$	1.069^{***}	0.12	1.008^{***}	0.044
ω_{α_h}	1.257^{***}	0.033	1.224^{***}	0.021
$\omega_{\alpha_{FT},\alpha_{PT}}$	33.064^{***}	6.947	32.58^{***}	5.925
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.743^{***}	0.149	0.942^{***}	0.064
$\omega_{\alpha_h,\alpha_{FT}}$	0.414	0.653	0.144	0.377
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	-0.537^{***}	0.099	-0.502^{***}	0.091
$\omega_{\alpha_h,\alpha_{PT}}$	0.078	0.528	0.16	0.373
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	-0.599^{***}	0.108	-0.483^{***}	0.094
Log-likelihood	20627.9		23652.8	
No. of observations	3121		3557	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table 13: Alternative full-time instruments and ever self-employed

	Past usual ret. age		Empl. health insurance	
	Coef	SE	Coef	SE
Health equation				
Full-time work utility	−0.031	0.022	−0.251*	0.135
Part-time work utility	0.038	0.024	0.279**	0.132
Age (minus 50)/10	0.201	0.154	0.121	0.198
Age deviation squared/100	−0.148***	0.039	−0.249***	0.052
Activities of daily living	−0.554***	0.023	−0.609***	0.031
Health conditions	−0.977***	0.035	−0.96***	0.043
Subsequent mortality	−0.783***	0.053	−0.703***	0.062
Full-time equation				
Health	0.442***	0.104	0.597***	0.078.
Constant	16.931***	1.753	14.933***	1.308
Age (minus 50)/10	−14.772***	1.62	−10.508***	1.116
Age deviation squared/100	1.625***	0.284	1.309***	0.249
Able to reduce hours	4.884***	0.544	3.796***	0.42
Past usual ret. age / Empl. health insurance	2.135***	0.199	−1.661***	0.197
Part-time equation				
Health	0.344***	0.111	0.505***	0.086
Constant	14.735***	1.704	12.886***	1.407
Age (minus 50)/10	−13.998***	1.599	−10.079***	1.124
Age deviation squared/100	2.045***	0.284	1.556***	0.239
Able to reduce hours	5.293***	0.564	4.28***	0.434
Past usual ret. age / Empl. health insurance	0.944***	0.182	−1.377***	0.204
Covariance parameters and cut-off estimates				
m_0	−4.367***	0.177	−4.702***	0.268
m_1.	−2.682***	0.174	−3.12***	0.291
m_2	−1.002***	0.174	−1.51***	0.319
m_3	0.786***	0.175	0.198	0.35
$\omega_{\alpha_{FT}}$	6.79***	0.721	5.077***	0.476
$\omega_{\alpha_{PT}}$	6.445***	0.722	4.842***	0.483
$\omega_{\varepsilon_{PT}}$	0.919***	0.116	0.914***	0.079
ω_{α_h}	1.252***	0.021	1.3***	0.041
$\omega_{\alpha_{FT},\alpha_{PT}}$	42.098***	9.419	23.88***	4.742
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.257	0.17	0.684***	0.147
$\omega_{\alpha_h,\alpha_{FT}}$	0.649	0.503	0.095	0.514
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	−0.377***	0.111	−0.413***	0.098
$\omega_{\alpha_h,\alpha_{PT}}$	0.616	0.487	−0.466	0.418
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	−0.319***	0.115	−0.54***	0.088
Log-likelihood	28835.6		23292.2	
No. of observations	4320		3461	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table 14: Full set of retirement eligibility age indicators and two part-time instruments

	Able to reduce hours		Ever self-employed		Union membership	
	Coef	SE	Coef	SE	Coef	SE
Health equation						
Full-time work utility	-0.188^{**}	0.084	-0.291^{***}	0.1	-0.14^{*}	0.081
Part-time work utility	0.197^{**}	0.079	0.318^{***}	0.099	0.162^{*}	0.088
Age (minus 50)/10	0.1	0.221	0.242	0.195	0.113	0.275
Age deviation squared/100	-0.209^{***}	0.053	-0.264^{***}	0.05	-0.164^{**}	0.066
Activities of daily living	-0.62^{***}	0.032	-0.567^{***}	0.03	-0.591^{***}	0.028
Health conditions	-0.949^{***}	0.045	-0.93^{***}	0.044	-0.99^{***}	0.042
Subsequent mortality	-0.691^{***}	0.07	-0.742^{***}	0.063	-0.758^{***}	0.064
Full-time equation						
Health	0.555^{***}	0.129	0.557^{***}	0.122	0.606^{***}	0.069
Constant	16.057^{***}	2.067	16.517^{***}	2.12	16.946^{***}	1.628
Age (minus 50)/10	-13.302^{***}	2.106	-11.569^{***}	1.778	-11.574^{***}	1.356
Age deviation squared/100	1.772^{***}	0.459	1.032^{***}	0.367	1.792^{***}	0.329
Between age 62 and normal ret. age	-1.071^{***}	0.223	-1.251^{***}	0.244	-0.879^{***}	0.196
Between normal ret. age and age 70	-1.256^{***}	0.324	-1.438^{***}	0.335	-0.952^{***}	0.278
Over age 70	-1.107^{**}	0.503	-1.13^{**}	0.512	-0.754^{*}	0.416
Between age 62 and normal ret. age (P)	-0.934^{***}	0.208	-0.809^{***}	0.189	-0.697^{***}	0.17
Between normal ret. age and age 70 (P)	-1.238^{***}	0.258	-0.923^{***}	0.24	-1.024^{***}	0.222
Over age 70 (P)	-1.946^{***}	0.387	-1.428^{***}	0.365	-1.335^{***}	0.311
Reduce hours / Self-employed/ Union-member	4.863^{***}	0.719	4.548^{***}	0.706	-2.107^{***}	0.316
Part-time equation						
Health	0.484^{***}	0.136	0.472^{***}	0.132	0.586^{***}	0.071
Constant	13.421^{***}	2.093	14.143^{***}	2.056	16.539^{***}	1.593
Age (minus 50)/10	-12.975^{***}	2.09	-11.162^{***}	1.754	-11.573^{***}	1.357
Age deviation squared/100	2.054^{***}	0.446	1.264^{***}	0.359	1.864^{***}	0.333
Between age 62 and normal ret. age	-0.711^{***}	0.227	-0.943^{***}	0.237	-0.806^{***}	0.198
Between normal ret. age and age 70	-0.733^{**}	0.327	-0.888^{***}	0.326	-0.829^{**}	0.282
Over age 70	-0.676	0.487	-0.704	0.485	-0.657	0.423
Between age 62 and normal ret. age (P)	-0.799^{***}	0.205	-0.596^{***}	0.182	-0.654^{***}	0.169
Between normal ret. age and age 70 (P)	-0.935^{***}	0.254	-0.542^{**}	0.236	-0.976^{**}	0.221
Over age 70 (P)	-1.38^{***}	0.361	-0.756^{**}	0.335	-1.197^{**}	0.304
Reduce hours / Self-employed/ Union-member	5.622^{***}	0.757	5.093^{***}	0.746	-2.172^{***}	0.322
Covariance parameters and cut-off estimates						
m_0	-4.698^{***}	0.31	-4.575^{***}	0.244	-4.667^{***}	0.341
m_1	-3.051^{***}	0.314	-2.962^{***}	0.255	-2.981^{***}	0.34
m_2	-1.338^{***}	0.321	-1.314^{***}	0.271	-1.26^{***}	0.34
m_3	0.471	0.329	0.435	0.29	0.566^{*}	0.341
$\omega_{\alpha_{FT}}$	6.153^{***}	0.783	6.617^{***}	0.898	5.518^{***}	0.582
$\omega_{\alpha_{PT}}$	5.794^{***}	0.781	6.305^{***}	0.888	5.511^{***}	0.581
$\omega_{\varepsilon_{PT}}$	0.822^{***}	0.105	0.945^{***}	0.118	0.999^{***}	0.021
ω_{α_h}	1.235^{***}	0.033	1.261^{***}	0.032	1.267^{***}	0.036
$\omega_{\alpha_{FT},\alpha_{PT}}$	34.673^{***}	9.269	40.543^{***}	11.38	30.353^{***}	6.397
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.529^{***}	0.182	0.558^{***}	0.195	0.983^{***}	0.027
$\omega_{\alpha_h,\alpha_{FT}}$	0.265	0.661	0.072	0.652	0.492	0.703
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	-0.496^{***}	0.14	-0.396^{***}	0.136	-0.659^{***}	0.072
$\omega_{\alpha_h,\alpha_{PT}}$	-0.05	0.544	-0.505	0.555	0.529	0.713
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	-0.513^{***}	0.134	-0.555^{***}	0.129	-0.643^{***}	0.075
Log-likelihood	16858.2		20833.7		20843.0	
No. of observations	2554		3146		3130	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. (P) stands for Married or unmarried partner. 3. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

5.2 Econometric model

The econometric model that we employ uses a simultaneous equations framework where the latent dependent variable of one equation enters as a covariate in the other equation(s). Furthermore, the error components of all equations are allowed to be correlated. This allows us to model the simultaneous nature of the relationship between health and the retirement decisions. Moreover, we exploit the panel nature of the HRS dataset, controlling for unobserved heterogeneity by allowing for random effects. Here, we analyze to which extent controlling for unobserved heterogeneity affects the baseline estimation results. To do this, we estimate a simultaneous equations model where we do not allow the error components to follow a random effects structure. Rather, we assume independence between the error components in different time periods, meaning that we do not control for unobserved individual heterogeneity. The estimation results are presented in Table 15. We find that when unobserved heterogeneity is not controlled for, the size, sign and significance of the coefficients change drastically, implying that individual specific unobserved effects play a significant role in the relationship between health and labor decisions.

An important feature of our model is that we distinguish between working full-time and part-time and model the two different types of market participation decisions in a multinomial probit framework. The main advantage of the multinomial probit model (MNP) over the commonly used multinomial logit model (MNL) is that it does not impose the very restrictive Independence of Irrelevant Alternatives (IIA) assumption (Train, 2009). However, a practical drawback of the MNP model can be that in some applications it is weakly identified (Keane, 1992). As the MNL can not be applied in our simultaneous equations framework, we cannot directly compare it against the MNP. Therefore, we will test the sensitivity of our model to imposing the IIA assumption by restricting the two underlying covariance parameters of the error components of the full-time and part-time utilities from equation (2) to both equal zero. This is done by restricting the observed covariance parameters of the error components of the utility difference of full-time work and retirement, and the utility difference of part-time work and retirement. This leads to a multinomial probit model that satisfies the IIA assumption. The resulting model is a nested version of our original model, meaning that we can formally test between the two models by means of a likelihood-ratio test. The likelihood-ratio test statistic takes a value of 847.4. The 1% critical value for the Chi-squared distribution with 2 degrees of freedom is $\chi^2_{2,0.99} = 9.21$ Hence, we reject the restricted model in favor of the unrestricted model at the 1% significance level. This finding provides evidence that allowing for correlation between the error terms of the equations (10) and (11) matters.

In order to assess how the instruments for labor and health help model identification, we estimate a model containing no instruments. The only variables included in this specification are the age and the squared age variables. We observe that without instrumenting the work decisions and health status, the log-likelihood value is significantly higher, meaning that including the instruments greatly improves model fit. In part, this happens because without imposing exclusion restrictions on the health and labor equations, we rely exclusively on the nonlinearity of the model to identify the parameters. We observe that the model without instrumental variables fails to identify full-time utility from part-time utility, implying that the age and squared age variables by themselves are not sufficient to distinguish between full-time and part-time work. As a result, multicollinearity arises in equation (7) as both full-time and part-time utility enter here, leading to unreliable standard errors.

Table 15: Results for model without Random Effects

	Baseline equation	
	Coef	SE
Health equation		
Full-time work utility	−0.068	0.064
Part-time work utility	0.002	0.014
Age (minus 50)/10	−0.111	0.08
Age deviation squared/100	−0.038	0.035
Activities of daily living	−0.601***	0.015
Health conditions	−0.793***	0.019
Subsequent mortality	−0.459***	0.015
Full-time equation		
Health	0.409***	0.019
Constant	2.747***	0.113
Age (minus 50)/10	−1.672***	0.181
Age deviation squared/100	0.171*	0.093
Between age 62 and normal ret. age	−0.147**	0.061
Between normal ret. age and age 70	−0.272***	0.099
Over age 70	−0.241	0.202
Between age 62 and normal ret. age (P)	−0.259***	0.04
Between normal ret. age and age 70 (P)	−0.351***	0.034
Over age 70 (P)	−0.624***	0.049
Part-time equation		
Health	0.885	0.833
Constant	−18.544	25.333
Age (minus 50)/10	3.801	7.123
Age deviation squared/100	−2.678	3.864
Between age 62 and normal ret. age	2.687	3.465
Between normal ret. age and age 70	3.645	4.718
Over age 70	3.792	5.547
Between age 62 and normal ret. age (P)	1.597	2.218
Between normal ret. age and age 70 (P)	1.24	1.854
Over age 70 (P)	2.1	2.888
Covariance parameters and cut-off estimates		
m_0	−3.141***	0.207
m_1	−2.101***	0.188
m_2	−1.042***	0.171
m_3	0.093	0.154
$\omega_{\varepsilon_{PT}}$	17.817	20.944
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	3.967	5.449
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	−0.228**	0.111
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	0.052	4.539
Log-likelihood	29577.3	
No. of observations	3519	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. (P) stands for Married or unmarried partner. 3. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

5.3 Control variables

In our baseline specification, we only control for a quadratic functional form of age. To test the sensitivity of the results to the inclusion of additional control variables, we estimate a model where we control for characteristics that might influence the labor market participation decision. We include the four control variables discussed in Section 3.3. These control variables are: a gender dummy indicating that the respondent is male, a dummy variable that indicates that the respondents has a college degree or a higher level of education, a variable for the size of the respondent's household and an indicator for blue collar work. The estimation results of this model are reported in Table 16. We find that the effect of gender on health is not significant in our model. The variable indicating that someone has a college degree or a higher level of education is positive and significant at the 1% level, indicating that education is beneficial to health, a common finding in the literature (Ross and Wu, 1995; Cutler and Lleras-Muney, 2006). The effect of household size on the health outcome is negative and significant at the 5% level. This could be due to the negative effect of household size on mental health as found by Gove and Geerken (1977), an aspect of health that we do not instrument. Blue-collar work is found to have a negative effect on health, significant at the 1% level, in line with the findings of Sickles and Taubman (1986). Gender is found to have a positive effect on the utility difference between full-time work and retirement and the utility difference between part-time and retirement, implying that, ceteris paribus, male respondents have a higher probability of working than female respondents. The effect of education on full-time work and part-time work is positive and significant at the 1% level, indicating that having a higher level of education increases the probability of working versus being retired. Household size has a positive effect on full-time and part-time work and is signifiant at the 5% level in both equations. For blue-collar work, no significant effects on labor market participation are found.

As discussed in Section 3.3, including relevant control variables in the model can affect the exogeneity of the instrumental variables if the control variables are correlated with one or more of the instruments. This is because including relevant control variables in the model means removing their effects from the error term. It can be expected that, through various mechanisms, some of the control variables are at least weakly correlated with the instruments of the labor market participation decisions and health status. Therefore, we perform the exclusion restrictions tests presented in Section 4.1 for the model including the control variables. Table 18 presents the results for the health instruments. In line with our expectations, we find that in this model, the Z-values are lower compared to Table 3. We fail to reject the instruments based on health conditions and physical functioning at the 10% significance level. Moreover, in this model we find stronger evidence against rejection of the null-hypothesis of exogeneity of the subsequent mortality instrument. Table 19 presents the Wald statistics for the joint significance of the set of retirement eligibility ages of the respondent and the set of retirement eligibility ages of the spouse in the health equation. In line with our expectations, the Wald statistics are lower compared to Table 4. These findings indicate that the evidence against rejection of the null-hypothesis of exogeneity of the health and labor instruments is stronger if we include control variables in the model.

Table 16: Results for model including control variables

	Baseline equation	
	Coef	SE
Health equation		
Full-time work utility	-0.103^{*}	0.058
Part-time work utility	0.134^{**}	0.055
Age (minus 50)/10	-0.121	0.476
Age deviation squared/100	-0.100^{**}	0.051
Gender	-0.077	0.111
Education	0.465^{***}	0.089
Household size	-0.034^{**}	0.017
Blue-collar work	-0.455^{***}	0.056
Activities of daily living	-0.569^{***}	0.03
Health conditions	-0.976^{***}	0.049
Subsequent mortality	-0.706^{***}	0.062
Full-time equation		
Health	0.729^{***}	0.052
Constant	12.001^{***}	1.435
Age (minus 50)/10	-8.37^{***}	1.231
Age deviation squared/100	0.857^{***}	0.305
Gender	0.696^{***}	0.249
Education	0.826^{***}	0.263
Household size	0.159^{**}	0.068
Blue-collar work	-0.035	0.242
Between age 62 and normal ret. age	-0.925^{***}	0.184
Between normal ret. age and age 70	-1.139^{***}	0.264
Over age 70	-0.817^{**}	0.391
Between age 62 and normal ret. age (P)	-0.526^{***}	0.152
Between normal ret. age and age 70 (P)	-0.697^{***}	0.208
Over age 70 (P)	-1.034^{***}	0.3
Part-time equation		
Health	0.71^{***}	0.056
Constant	11.712^{***}	1.47
Age (minus 50)/10	-8.318^{***}	1.234
Age deviation squared/100	0.937^{***}	0.298
Gender	0.518^{**}	0.246
Education	0.844^{***}	0.261
Household size	0.148^{**}	0.068
Blue-collar work	-0.031	0.135
Between age 62 and normal ret. age	-0.853^{***}	0.192
Between normal ret. age and age 70	-1.025^{***}	0.275
Over age 70	-0.727^{*}	0.397
Between age 62 and normal ret. age (P)	-0.517^{***}	0.152
Between normal ret. age and age 70 (P)	-0.704^{***}	0.207
Over age 70 (P)	-0.999^{***}	0.297

- continued from previous page

Covariance parameters and cut-off estimates		
m_0	-4.771^{***}	0.493
m_1	-3.146^{***}	0.498
m_2	-1.447^{***}	0.503
m_3	0.371	0.509
$\omega_{\alpha_{FT}}$	5.434^{***}	0.612
$\omega_{\alpha_{PT}}$	5.369^{***}	0.614
$\omega_{\varepsilon_{PT}}$	0.97^{***}	0.029
ω_{α_h}	1.244^{***}	0.085
$\omega_{\alpha_{FT},\alpha_{PT}}$	29.134^{***}	6.615
$\omega_{\varepsilon_{FT},\varepsilon_{PT}}$	0.957^{***}	0.047
$\omega_{\alpha_h,\alpha_{FT}}$	0.726	1.454
$\omega_{\varepsilon_h,\varepsilon_{FT}}$	-0.789^{***}	0.058
$\omega_{\alpha_h,\alpha_{PT}}$	0.749	1.484
$\omega_{\varepsilon_h,\varepsilon_{PT}}$	-0.768^{***}	0.064
Log-likelihood	22516.6	
No. of observations	3461	

Notes: 1.Full-time and part-time work utility are differenced with respect to retirement utility 2. (P) stands for Married or unmarried partner. 3. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table 18: Z-test scores for exogeneity of health instruments

	Health conditions		Physical functioning		Subsequent mortality	
	Full-time	Part-time	Full-time	Part-time	Full-time	Part-time
Z-statistic	-1.11	-1.26	-1.61	-0.93	-1.70	-1.57
$p \geq Z$	0.27	0.21	0.11	0.35	0.09	0.12

1. Standard errors used to calculate the Z-statistics are derived using the Newey-West sandwich estimator (Newey and West, 1986) and are robust to heteroskedasticity and autocorrelation. 2. The estimated model includes the set of control variables discussed in the data section.

Table 19: Wald-statistics for scores for exogeneity of labor instruments

	Health equation	
	Retirement eligibility ages	Retirement eligibility ages (P)
Wald statistic	1.26	1.50
$p \geq Wald$	0.74	0.68

1. Standard errors used to calculate the Z-statistics are derived using the Newey-West sandwich estimator (Newey and West, 1986) and are robust to heteroskedasticity and autocorrelation. 2. (P) stands for married or unmarried partner 3. The estimated model includes the set of control variables discussed in the data section.

6 Conclusion

This paper examines the relationship between health, full-time work and part-time-work, using the HRS dataset. We control for endogeneity resulting from simultaneity by estimating a simultaneous equations model using the full-information maximum likelihood method. By allowing the error terms of the health and the labor equations to be correlated, we control for another source of endogeneity resulting from unobserved factors that affect both the work decisions and health status.

We exploit the panel data nature of the HRS and control for unobserved heterogeneity by allowing a random effects structure for the error terms and find that this greatly improves the efficiency of the estimates compared to when the error components are assumed to be independent over time.

Moreover, we model the discrete nature of the dependent variables by employing an ordered probit model for self-reported health and a multinomial probit model for the labor market participation decisions. Allowing for correlation between the error components of the equations for full-time work and part-time work results in significantly more efficient estimates than if the parameters measuring the covariance between the error components of the equations are restricted.

The estimation results confirm the common finding in the literature that health has a positive effect on the probability of working. For the reverse effect, we find that the effect of working full-time on health is negative and that the effect of working part-time on health is positive, although this finding is not always robust in alternative specifications.

Furthermore, the test for the joint significance of the parameter estimates measuring the different potential sources of endogeneity rejects the hypothesis that health and labor market participation are exogenous. The signs of the estimates imply that models that do not control for endogeneity might underestimate the effects of health on full-time work. Therefore, with regards to the aging population in particular, controlling for endogeneity in the relationship between health and labor is of vital importance.

As marginal effects cannot be estimated for this model, we assess the magnitude of the health and labor effects by prediction. For policy makers, understanding the sizes of such effects is important, because they can be used, for example, to estimate the indirect costs of health problems to society through their effects on the labor supply in older age.

There are a number of future research directions that can be pursued. First, we consider three health instruments based on health conditions, physical functioning and subsequent mortality. In order to improve upon this instrument set, one could consider lagged versions of health instruments. Cai (2010) estimates a specification where he uses lagged indicators of being a smoker, being a heavy drinker, having physical functioning limitations, or having health conditions, while he uses the values of those variables for the current year as control variables. Validity of this specification is based on the assumption that the labor market is flexible so that indicators of past health are not relevant if we control for current values of those health indicators.

Second, we only include a limited number of control variables. The main reason for this is that many control variables are potentially endogenous to health or labor. An example of such a variable is income. This variable is highly likely to be endogenous in the labor equation given that an increased number of working hours increases income. A possible strategy would be to estimate the model for different groups of individuals with different income levels and compare the differences. This strategy also allows for heterogeneity in the treatment effects. Cai (2010) finds that the effect of labor force participation on health is different for males and females. Therefore, it can be interesting to estimate our model for males and females separately, allowing

the effects to be heterogeneous between males and females.

Third, we do not use a formal test for overidentifying restrictions such as the test proposed by Lee (1992). The results of the test that we use can be unreliable if not all instruments are exogenous. Therefore, in order to obtain more conclusive evidence about whether or not the imposed exclusion restrictions are valid, a test along the lines of Lee (1992) might be conducted.

Finally, in order to identify the effects of full-time and part-time work on health, it is important that the instrumental variables that we employ to instrument these endogenous regressors offer independent sources of exogenous variation for both regressors. Sanderson and Windmeijer (2016) provide a method to formally test whether an endogenous regressor is weakly identified in a linear model with multiple endogenous regressors. A formal test along the lines of Sanderson and Windmeijer (2016) for our nonlinear model could be used to test whether a set of instruments provides independent sources of exogenous variation for both endogenous regressors.

References

D. Aaronson and E. French. The effect of part-time work on wages: evidence from the social security rules. *Journal of Labor Economics*, 22(2):329–252, 2004.

T. Amemiya. The estimation of a simultaneous equation generalized probit model. *Econometrica: Journal of the Econometric Society*, pages 1193–1205, 1978.

K. H. Anderson and R. V. Burkhauser. The importance of the measure of health in empirical estimates of the labor supply of older men. *Economics letters*, 16(3-4):375–380, 1984.

K. H. Anderson and R. V. Burkhauser. The retirement-health nexus: a new measure of an old puzzle. *Journal of Human Resources*, pages 315–330, 1985.

G. J. Bazzoli. The early retirement decision: new empirical evidence on the influence of health. *Journal of human resources*, pages 214–234, 1985.

G. M. Becker, M. H. DeGroot, and J. Marschak. Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232, 1964.

S. Behncke. Does retirement trigger ill health? *Health economics*, 21(3):282–300, 2012.

A. Biró. Outpatient visits after retirement in Europe and the US. *International Journal of Health Economics and Management*, 16(4):363–385, 2016.

D. M. Blau. Labor force dynamics of older married couples. *Journal of Labor Economics*, 16 (3):595–629, 1998.

E. Bonsang, S. Adam, and S. Perelman. Does retirement affect cognitive functioning? *Journal of Health Economics*, 31(3):490–501, 2012.

A. Borsch-Supan, V. Hajivassiliou, and L. J. Kotlikoff. Health, children, and elderly living arrangements: A multiperiod-multinomial probit model with unobserved heterogeneity and autocorrelated errors. In *Topics in the Economics of Aging*, pages 79–108. University of Chicago Press, 1992.

J. Bound. Self-reported vs. objective measures of health in retirement models. Technical report, National Bureau of Economic Research, 1989.

J. Bound. Self-reported versus objective measures of health in retirement models. *The Journal of Human Resources*, 26(1):106–138, 1991.

J. Bound and T. Waidmann. Estimating the health effects of retirement. 2007.

J. Bound, D. A. Jaeger, and R. M. Baker. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American statistical association*, 90(430):443–450, 1995.

J. Bound, M. Schoenbaum, T. R. Stinebrickner, and T. Waidmann. The dynamic effects of health on the labor force transitions of older workers. *Labour Economics*, 6(2):179–202, 1999.

C. Brown. The role of conventional retirement age in retirement decisions (july 2006). *Michigan Retirement Research Center Research Paper No. WP 2006-120*, 2006.

J. S. Butler and R. Moffitt. A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica: Journal of the Econometric Society*, pages 761–764, 1982.

K. E. Cahill, M. D. Giandrea, and J. F. Quinn. Retirement patterns and the macroeconomy, 1992–2010: the prevalence and determinants of bridge jobs, phased retirement, and reentry among three recent cohorts of older Americans. *The Gerontologist*, 55(3):384–403, 2015.

L. Cai. The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics*, 17(1):77–90, 2010.

L. Cai and G. Kalb. Health status and labour force participation: evidence from australia. *Health economics*, 15(3):241–261, 2006.

L. Cai, G. Kalb, et al. Health status and labour force status of older working-age australian men. *Australian Journal of Labour Economics*, 10(4):227, 2007.

M. Campolieti. Disability and the labor force participation of older men in canada. *Labour economics*, 9(3):405–432, 2002.

K. K. Charles. Is retirement depressing? Labor force inactivity and psychological well-being in later life. *Research in Labor Economics*, 23:269–299, 2004.

T. N. Chirikos. The relationship between health and labor market status. *Annual Review of Public Health*, 14(1):293–312, 1993.

N. B. Coe and G. Zamarro. Retirement effects on health in Europe. *Journal of Health Economics*, 30(1):77–86, 2011.

D. M. Cutler and A. Lleras-Muney. Education and health: evaluating theories and evidence. Technical report, National bureau of economic research, 2006.

C. F. Daganzo, F. Bouthelier, and Y. Sheffi. Multinomial probit and qualitative choice: A computationally efficient algorithm. *Transportation Science*, 11(4):338–358, 1977.

D. Dave, I. Rashad, and J. Spasojevic. The effects of retirement on physical and mental health outcomes. *Southern Economic Journal*, 75(2):497–523, 2008.

P. A. Diamond and J. A. Hausman. Individual retirement and savings behavior. *Journal of Public Economics*, 23(1-2):81–114, 1984.

R. Disney, C. Emmerson, and M. Wakefield. Ill health and retirement in britain: A panel data-based analysis. *Journal of health economics*, 25(4):621–649, 2006.

D. S. Dwyer and O. S. Mitchell. Health problems as determinants of retirement: Are self-rated measures endogenous? *Journal of health economics*, 18(2):173–193, 1999.

R. G. Ehrenberg, P. Rosenberg, and J. Li. Part-time employment in the united states. In R. A. Hart, editor, *Employment, unemployment and labor utilization*, pages 256–281. Unwin Hyman, Inc., 1988.

P. Eibich. Understanding the effect of retirement on health: mechanisms and heterogeneity. *Journal of Health Economics*, 43:1–12, 2015.

C. Eisdorfer and F. Wilkie. Stress, disease, aging and behavior. *Handbook of the psychology of aging*, pages 251–275, 1977.

D. Ekerdt, S. DeViney, and K. Kosloski. Profiling plans for retirement. *Journal of Gerontology*, 51B(3):S140–S149, 1996.

D. J. Ekerdt, L. Baden, R. Bosse, and E. Dibbs. The effect of retirement on physical health. *American journal of public health*, 73(7):779–783, 1983.

S. L. Ettner. New evidence on the relationship between income and health. *Journal of health economics*, 15(1):67–85, 1996.

M. K. Forbes, K. M. Spence, V. M. Wuthrich, and R. M. Rapee. Mental health and wellbeing of older workers in Australia. *Work, Aging and Retirement*, 1(2):202–213, 2015.

E. French and J. B. Jones. The effects of health insurance and self-insurance on retirement behavior. *Econometrica*, 79(3):693–732, 2011.

E. French and J. B. Jones. Health, health insurance, and retirement: A survey. *Annual Review of Economics*, 9:383–409, 2017.

J. Geweke. Antithetic acceleration of monte carlo integration in bayesian inference. *Journal of Econometrics*, 38(1-2):73–89, 1988.

J. Geweke, M. P. Keane, D. E. Runkle, et al. *Statistical Inference in the Multinominal Multiperiod Probit Model*. Federal Reserve Bank of Minneapolis, Research Department, 1994.

W. R. Gove and M. R. Geerken. The effect of children and employment on the mental health of married men and women. *Social Forces*, 56(1):66–76, 1977.

M. Grossman. On the concept of health capital and the demand for health. *Journal of Political economy*, 80(2):223–255, 1972.

M. Grossman. The human capital model of the demand for health. Technical report, National Bureau of Economic Research, 1999.

A. L. Gustman and T. L. Steinmeier. Modeling the retirement process for policy evaluation and research. *Monthly Labor Review*, 107(7):26–33, 1984a.

A. L. Gustman and T. L. Steinmeier. Partial retirement and the analysis of retirement behavior. *Industrial and Labor Relations Review*, 37(3):403–415, 1984b.

A. L. Gustman and T. L. Steinmeier. Retirement in dual-career families: A structural model. *Journal of Labor Economics*, 18(3):503–545, 2000a.

A. L. Gustman and T. L. Steinmeier. Retirement outcomes in the health and retirement study. *Social Security Bulletin*, 63(4):57–71, 2000b.

A. L. Gustman and T. L. Steinmeier. Minimum hours constraints, job requirements and retirement. *NBER Working Paper 10876*, 2004.

A. L. Gustman and T. L. Steinmeier. Integrating retirement models: understanding household retirement decisions. *Research in Labor Economics*, 40:79–112, 2014.

P. Haan and M. Myck. Dynamics of health and labor market risks. *Journal of Health Economics*, 28(6):1116–1125, 2009.

V. Hajivassiliou. Some practical issues in maximum simulated likelihood. *Simulation-Based inference in econometrics: Methods and Applications*, pages 71–99, 2000.

V. Hajivassiliou, D. McFadden, and P. Ruud. Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results. *Journal of econometrics*, 72(1-2):85–134, 1996.

V. A. Hajivassiliou. *Simulation estimation methods for limited dependent variable models*. Yale University, 1992.

V. A. Hajivassiliou and P. A. Ruud. Classical estimation methods for ldv models using simulation. 1993.

D. Hallberg, P. Johansson, and M. Josephson. Is an early retirement offer good for your health? Quasi-experimental evidence from the army. *Journal of Health Economics*, 44:274–285, 2015.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

J. J. Heckman. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. 1981.

J. C. Henretta, A. M. O'Rand, and C. G. Chan. Joint role investments and synchronization of retirement: A sequential approach to couples' retirement timing. *Social Forces*, 71(4): 981–1000, 1993.

M. Honig and G. Hanoch. Partial retirement as a separate mode of retirement behavior. *The Journal of Human Resources*, 20(1):21–46, 1985.

J. L. Horowitz, J. M. Sparmann, and C. F. Daganzo. An investigation of the accuracy of the clark approximation for the multinomial probit model. *Transportation Science*, 16(3): 382–401, 1982.

D. R. Hyslop. State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, 67(6):1255–1294, 1999.

E. L. Idler and S. V. Kasl. Self-ratings of health: do they also predict change in functional ability? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 50(6):S344–S353, 1995.

G. W. Imbens and J. D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

M. Insler. The health consequences of retirement. *Journal of Human Resources*, 49(1):195–233, 2014.

D. W. Johnston and W.-S. Lee. Retiring to the good life? the short-term effects of retirement on health. *Economics Letters*, 103(1):8–11, 2009.

M. P. Keane. A note on identification in the multinomial probit model. *Journal of Business & Economic Statistics*, 10(2):193–200, 1992.

M. P. Keane. A computationally practical simulation estimator for panel data. *Econometrica: Journal of the Econometric Society*, pages 95–116, 1994.

D. S. Kenkel. Should you eat breakfast? estimates from health production functions. *Health economics*, 4(1):15–29, 1995.

M. Kerkhofs, M. Lindeboom, and J. Theeuwes. Retirement, financial incentives and health. *Labour Economics*, 6(2):203–227, 1999.

H. Kim and S. DeVaney. The selection of partial or full retirement by older workers. *Journal of Family and Economic Issues*, 26(3):371–394, 2005.

L.-F. Lee. Amemiya‘s generalized least squares and tests of overidentification in simultaneous equation models with qualitative or limited dependent variables. *Econometric Reviews*, 11 (3):319–328, 1992.

O. Lundberg and K. Manderbacka. Assessing reliability of a measure of self-rated health. *Scandinavian journal of social medicine*, 24(3):218–224, 1996.

A. Macbride. Retirement as a life crisis: Myth or reality? a review. *Canadian Psychiatric Association Journal*, 21(8):547–556, 1976.

F. Mazzonna and F. Peracchi. Ageing, cognitive abilities and retirement. *European Economic Review*, 56(4):691–710, 2012.

F. Mazzonna and F. Peracchi. Unhealthy retirement? *Journal of Human Resources*, page Forthcoming, 2016.

R. D. McKelvey and W. Zavoina. A statistical model for the analysis of ordinal level dependent variables. *Journal of mathematical sociology*, 4(1):103–120, 1975.

K. Neuman. Quit your job and get healthier? The effect of retirement on health. *Journal of Labor Research*, 29(2):177–201, 2008.

W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix, 1986.

M. A. Okun, W. A. Stock, M. J. Haring, and R. A. Witter. Health and subjective well-being: A meta-analyis. *The International journal of aging and human development*, 19(2):111–132, 1984.

D. O. Parsons. The male labour force participation decision: health, reported health, and economic incentives. *Economica*, 49(193):81–91, 1982.

S. Rohwedder and R. J. Willis. Mental retirement. *Journal of Economic Perspectives*, 24 (1): 119–138, 2010.

C. E. Ross and C.-l. Wu. The links between education and health. *American sociological review*, pages 719–745, 1995.

C. Ruhm. Bridge jobs and partial retirement. *Journal of Labor Economics*, 8(4):482–501, 1990.

J. Rust and C. Phelan. How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica*, 65(4):781–832, 1997.

E. Sanderson and F. Windmeijer. A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221, 2016.

J. D. Sargan. The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the Econometric Society*, pages 393–415, 1958.

R. C. Sickles and P. Taubman. An analysis of the health and retirement status of the elderly. *Econometrica: Journal of the Econometric Society*, pages 1339–1356, 1986.

S. Stern. Measuring the effect of disability on labor force participation. *Journal of human Resources*, pages 361–395, 1989.

K. Train. *Discrete Choice Methods with Simulation.* Cambridge University Press, 2009.

D. D. Vaus, Y. Wells, H. Kendig, and S. Quine. Does gradual retirement have better outcomes than abrupt retirement? Results from an Australian panel study. *Ageing & Society*, 27(5): 667–682, 2007.

H. Westerlund, M. Kivimäki, A. Singh-Manoux, M. Melchior, J. E. Ferrie, J. Pentti, M. Jokela, C. Leineweber, M. Goldberg, M. Zins, et al. Self-rated health before and after retirement in france (gazel): a cohort study. *The Lancet*, 374(9705):1889–1896, 2009.

J. Wilde. Identification of multiple equation probit models with endogenous dummy regressors. *Economics letters*, 69(3):309–312, 2000.