Pensions, retirement, and the financial position of the elderly
Been, J.

Author: Been, Jim
Title: Pensions, retirement, and the financial position of the elderly
Issue Date: 2015-04-30

4 Estimating a panel data sample selection model with part-time employment: Selection issues in wages over the life-cycle

Abstract

This paper proposes a new panel data sample selection model for estimating wages over the life-cycle. The new estimator is an extension of the work of Rochina-Barrachina (1999) who proposed an estimator for panel data selection models where both the selection and the wage equation contain individual effects allowed to be correlated with the observable variables. Instead of solely correcting for systematic differences between those who work and those who do not work (*binary* selection), we extend the model by taking into account part-time and full-time work (*ordered* selection). Since part-time employment decisions provide additional information about unobserved characteristics. Our proposed method is likely to estimate improved wage profiles compared to models that use a binary selection indicator. The newly proposed estimator is applied to a large administrative data set based on Dutch tax records (2001-2011). The application allows us to analyze selection effects in part-time and full-time

This chapter is co-authored by Marike Knoef. We gratefully acknowledge Netspar and Instituut Gak for their financial contribution to this project. We thank all participants of the Netspar Workshop on Pensions, Retirement, and the Financial Position of the Elderly, the CPB Netherlands Bureau for Economic Policy Analysis Seminar and the conference participants at the International Association for Applied Econometrics 2014, London, the Econometric Society European Meetings 2014, Toulouse as well as the Netspar Pension Day 2014, Utrecht. More specifically, we would like to thank Rob Alessie, Jonneke Bolhaar, Jan Bonenkamp, Lans Bovenberg, Koen Caminada, Bart Cockx, Anja Deelen, Egbert Jongen, Stefan Hochguertel, Adriaan Kalwij, Stéphanie Payet, Peter van Santen, Ola Vestad, Daniel van Vuuren, Bas ter Weel, Mathijn Wilkens, Jeffrey Wooldridge and Bram Wouterse for providing us with valuable comments at different stages of the paper.

employment as well as the part-time wage penalty over the life-cycle. Education-specific life-cycle wage profiles show the existence of positive selection. For the average man, we do not find a part-time wage penalty. For the average low- and high-educated woman, we find part-time wage penalties of about 30%.

4.1 Introduction

Aging of the population confronts society with a growing number of dilemmas regarding the sustainability of public finances and collective arrangements. In OECD countries, pension systems are affected by demographic aging (OECD 2011b) and many countries have implemented or proposed reforms to alleviate the stress on the sustainability of the system primarily by increasing the statutory retirement age, making pension benefits less generous and increasing contribution rates. Forecasting the resources that people have available for post-retirement consumption is crucial when evaluating the impact of these reforms on government finances and financial well-being of retirees.

Since most pension formulas for pension accumulation are based on earnings during working life, life-cycle wage profiles are crucial in determining income available at retirement. Wages and wage processes are therefore a central component in life-cycle models. Especially, wage uncertainty and the persistence of income shocks play an important role in life-cycle models of consumption- and savings behavior that are used to evaluate retirement savings adequacy (Scholz et al. 2006). A life-cycle earnings model can also be used to simulate future (occupational) pension accumulation and the consequences of proposed reforms for such future pension accumulation (Borella 2004). A good understanding of the life-cycle wage profile is vital in this literature because deviations from the estimated deterministic component of the life-cycle wage profile are supposed to be the result of shocks.

Life-cycle models can be used to analyze retirement savings adequacy (Scholz et al. 2006). The conclusions of such analyses depend on the correct specification of the life-cycle wage-profile. However, the wage profile estimated in life-cycle models in general does not consider selection effects into work. Wages are likely to be observed non-randomly over the life-cycle, e.g. wages are only observed for people who are working. These same individuals may earn a different potential wage than the individuals who are not working. Neglecting this non-random selection into work may bias estimated wages (Heckman 1979) and wage-profiles (Casanova 2013).

The main objective of this paper is to estimate life-cycle wage profiles of persons in wage employment. We do not distinguish other labor market statuses like self-employment, unemployment, disability, early retirement and other inactivity.[1] We estimate life-cycle wage profiles by using panel data sample selection models with special attention given to selection into full-time and part-time employment. The incorporation of part-time employment is important as part-time employment plays an important role throughout the life-cycle for both men and women. Women tend to prefer part-time employment jobs in general because of the possibility to combine work and care (Booth and Van Ours 2008, Gregory and Connolly 2008). Such part-time employment is often associated with a lower wage than full-time wages among women (Manning and Petrongolo 2008). Among men, part-time employment is often preferred at older ages (Kantarci and Van Soest 2008) as a way to reduce working hours prior to full retirement (e.g. Ruhm (2006), Cahill et al. (2006)). Such end-of-career transitions often imply substantial drops in wages (Aaronson and French 2004, Casanova 2013, Hurd 1996, Johnson and Neumark 1996).

To estimate life-cycle wage profiles using a panel data sample selection model with part-time employment, we propose a new estimator that extends the work of Rochina-Barrachina (1999). Rochina-Barrachina (1999) proposed an estimator for panel data selection models where both the selection and the wage equation contain individual effects allowed to be correlated with the observable variables.[2] Compared to Rochina-Barrachina

[1] Introducing self-employment as a separate state would include another endogenous decision. Also, we do not have information on the number of hours worked by the self-employed.

[2] Other studies dealing with the estimation of panel data sample selection models are Wooldridge (1995) and Kyriazidou (1997). Dustmann and Rochina-Barrachina (2007) provide a comparison of these three aforementioned estimation methods. Endogeneity

(1999), who uses a binary selection rule in the selection equation, we implement an ordered selection rule. By using an ordered indicator instead of a binary selection indicator we are able to take into account extra information regarding unobserved individual characteristics, such as ability and preferences, from selection into part-time and full-time work that may influence wages. Instead of only correcting for systematic differences between those who work and those who do not work, we also take into account unobserved differences between those who are employed part-time and full-time in a panel data sample selection model.

Like Rochina-Barrachina (1999) we eliminate individual specific effects from the equation of interest by taking first- and higher order differences. Furthermore, a conditional mean independence assumption (Wooldridge 1995) is made to deal with the possible correlations between the unobserved individual specific effects and the explanatory variables in the selection equation. In the literature, discrete choice models have been used to analyze part-time and full-time wages, amongst others, by Ermisch and Wright (1993), Dustmann and Schmidt (2000). In contrast to these papers we use a combination of a bivariate ordered probit selection model and a wage equation in differences in order to eliminate individual specific unobserved effects nonparametrically in the second stage. The advantage of using differences in the wage equation is that it allows for an unknown conditional mean of the individual effects.

To estimate the model we use administrative data from the Dutch tax office for the years 2001-2011, which are more representative and reliable than survey data which are often used for the estimation of wage profiles.[3] Our proposed estimator allows us to analyze selection effects, selection into full-time and part-time employment, the part-time wage penalty and the effect of career breaks on wages over the life-cycle.

Earlier contributions to selection into work over the life-cycle shows a diverse picture. Ejrnaes and Kunze (2011) show the existence of negative selection in reentering full-time work after birth among German women.

issues and dynamic panel data sample selection models are dealt with in Semykina and Wooldridge (2010) and Semykina and Wooldridge (2011) respectively.

[3]Most studies analyzing life-cycle wage profiles rely on survey data from PSID. A number of shortcomings of the PSID for analyzing earnings dynamics are mentioned in Pischke (1995).

However, whether selection is positive or negative is found to possibly change over time among women (Mulligan and Rubinstein 2008). Myck (2010) finds that British men approaching the retirement age and who maintain their employment status are more likely to be the lower wage individuals (e.g. negative selection), whereas German men with higher wages are more likely to remain employed (e.g. positive selection). For the US, Casanova (2010) finds negative selection for older men. Using different selection terms Casanova (2013) does not find evidence for selection effects among older men (50+) at all.

For men, the results of applying our two-step estimator suggest the existence of positive selection into work over the life-cycle. This is in contrast with the results we obtain when using the binary selection correction proposed by Rochina-Barrachina (1999). Applying a binary selection indicator suggests negative selection into work. However, adding extra information using an ordered selection indicator changes the sign of selection. We also find positive selection into part-time employment and full-time employment among both men and women as well among low-educated and high-educated groups. Actual selection corrected life-cycle wage profiles however differ between these groups. Estimating education-specific models, we find no part-time wage penalties for the average low- and high educated man respectively. For the average woman, we find part-time wage penalties of 30% and 34% for low- and high-educated women respectively. This wage differential between part-time and full-time work may be a compensation for the ability to combine work with care and a consequence of less experience being accumulated (Boeri and Van Ours 2008). Career breaks have a significant downward effect on life-cycle wages for both men and women although the effect is somewhat more pronounced among men.

The proposed two-step estimator in this paper is likely to be useful in all applications of life-cycle earnings models as the second-stage wage equation is likely to give better estimates of the coefficients of wages over the life-cycle than wage profiles estimated without correction for selection[4] or with binary selection correction.[5] Applications of the model can vary from estimating life-cycle models (Gourinchas and Parker 2002, Scholz et al. 2006), analyzing earnings inequality (Baker and Solon 2003, Cappellari 2004, Haider 2001) to microsimulation exercises (Borella 2004).

The paper proceeds as follows. First, we describe the administrative data, the selection of the sample, and we provide a descriptive analysis of observed full-time and part-time wages over the life-cycle for men and women (section 4.2). Second, section 4.3 describes the basic model and explains the empirical specification. Section 4.4 shows the estimation results. Education-specific estimates are shown in section 4.4.3. Finally, section 4.5 concludes to what extent it is important to correct life-cycle wage profiles for selection into work and hours.

4.2 Data

The data in this study are taken from the 2001-2011 Income Panel Study from the Netherlands (IPO, CBS 2009a), the 2001-2011 Data on working hours (Baanprsjaarbedragtab, CBS 2010a) and the 2001-2011 data on the highest level of education (Hoogsteopltab, CBS 2010b). All three data sets are gathered by Statistics Netherlands. The IPO, a representative sample from the Dutch population, consists of an administrative panel dataset of, on average, 95,000 selected individuals per year who are followed longitudinally. Sampling is based on individuals' national security number, and the selected individuals are followed for as long as they are residing in the Netherlands on December 31 of the sample year. Individuals born

[4] To bypass possible selection papers focused on prime-aged males who are generally assumed to work to estimate wage profiles. MaCurdy (1982), Abowd and Card (1989), Baker (1997), Lillard and Reville (1999), Haider (2001), Meghir and Pistaferri (2004), Heathcote et al. (2010), Storesletten et al. (2004), Moffitt and Gottschalk (2012), Guvenen (2009), Altonji et al. (2009), Gottschalk and Zhang (2010), Ziliak et al. (2011) and Moffitt and Gottschalk (2011) (US). Dickens (2000), Ramos (2003) and Kalwij and Alessie (2007) (UK). Cappellari (2004) and Borella (2004) (Italy). Baker and Solon (2003) (Canada). Bonke et al. (2011) (Germany). Santos and Souza (2007) (Brazil). Magnac et al. (2011) (France). Sologon and O'Donoghue (2009) (Europe). As a consequence, the results of these models can not be generalized to women and persons approaching the retirement age (Kassi 2013).

[5] Such as Casanova (2010, 2013), Ejrnaes and Kunze (2011), Hanoch and Honig (1985), Johnson and Neumark (1996), Mulligan and Rubinstein (2008), Myck (2010).

in the Netherlands enter the panel for the first time in the year of their birth, and immigrants to the Netherlands in the year of their arrival. The main advantages of using this administrative dataset compared to using survey data for our analysis are, the large sample size, the long panel aspect of the data, the accuracy of tax data compared to survey questions, and representativeness. *Baanprsjaarbedragtab* contains information about working hours (the number of hours worked in proportion to a yearly full time job) for the whole Dutch population. *Hoogsteopltab* provides information on the highest level of education for a subsample of the Dutch population. We merge this information with the internationally standardized *ISCED3* measures of educational levels. The three data sets are merged based on the individuals' personal identifier.

Variable definitions and data selection 4.2.1

The dependent variables in our analysis is the real full-time equivalent wage expressed in 2010 euros. To construct wages, we divide yearly earnings by the proportion of hours worked relative to a full time job. This leaves us with a yearly full-time equivalent wage. Inevitably, we do not observe wages for people that do not work.

In this study we select individuals between the ages of 24 and 64 (309,025 observations for men and 305,678 observations for women). In the estimates, we only use information of persons born no later than 1980. Disentangling age, period and cohort effects works better when an individual is observed over a long time-span. Persons born later than 1980 are only available in the years 2006-2011 at relatively young ages.

Subsequently, we drop some outliers. First, we drop persons who worked less than one-twelfth of a full-time year. We argue they work to little to calculate a reliable wage. Second, we drop observations where the wage rate is higher than the 99th-percentile[6]. In this way we correct for possible measurement error in either earnings or the full-time employment factor leading to a very high wage. Third, we delete observations where the wage is lower than the minimum wage since the minimum wage is

[6]149,681 euros for men and 89,930 euros for women on average.

legally binding (we take into account yearly differences in the minimum wage level). Fourth, observations are dropped if a year-to-year-change in the wage rate is lower than -50% or higher than 80%. It is highly unlikely that these persons face a year-to-year change in their wage that is due to promotion or demotion. It is more likely that such big changes in year-to-year wages are a consequence of measurement error in the part-time employment factor. Finally, since a lot of people retire during the year observations about the last year of work before retirement are sensitive to mistakes in the number of hours worked in that year. Therefore, we drop observations for which the wage rate dropped more than 30% or increased more than 80% in the last year before retirement.

For the analysis that differentiates between education levels, we end up with 87,401 men and 84,757 women for whom the education level is known. We use population weights to make the sample representative with respect to age, gender, marital status, province, household size and the age of the head of the household.

4.2.2 Descriptive statistics

Table 4.1 shows the development of earnings and wages in our period of observation (2001-2011). The table shows that labor income (including zeros for non-workers) is over time for men and increased for women. Also the average and median wages rates (second column) are slightly increasing over time for men and women. For women, we observe that the average part-time employment factor (which is equal to one if full-time employed throughout the year) increased substantially over the years 2001-2011 from 0.39 to 0.47. For men, the table indicates that average wages are quite stable over time while median wages seem to have increased over time.

Table 4.1 solely focuses on trends over time. To gain insight in wage-differences over the life-cycle and between cohorts we construct age-cohort figures. Figure 4.1 presents average earnings for men and women (including those who do not work). For men, average earnings are about 20,000 euros per year at the age of 25 and grow up to about 35,000 euros per year around the age of 50. After the age of 50, we observe a decline

Table 4.1: Descriptives of real earnings and wage rates

Year	Average earnings[a]	Average wage	Median wage	S.D.	Part-time factor	Obs.[b]
Men						
2001	30,266	43,212	37,948	19,249	0.72	26,142
2002	30,047	43,160	37,974	19,025	0.72	25,764
2003	29,947	43,793	38,511	19,402	0.71	25,891
2004	29,933	44,311	39,076	19,883	0.70	25,717
2005	29,710	44,357	39,005	20,140	0.70	25,686
2006	29,856	44,474	39,196	20,153	0.70	25,823
2007	29,978	44,199	38,906	19,989	0.70	25,954
2008	30,213	44,299	39,062	20,097	0.71	25,820
2009	30,196	44,973	39,752	20,300	0.70	25,913
2010	29,703	44,913	39,692	20,479	0.70	25,831
2011	29,813	44,774	39,376	20,928	0.70	25,552
Women						
2001	13,007	33,564	31,149	11,879	0.39	24,118
2002	13,398	33,908	31,574	11,987	0.40	23,926
2003	13,573	34,265	31,911	12,008	0.40	24,177
2004	13,745	34,634	32,198	12,309	0.40	24,127
2005	13,900	34,688	32,003	12,749	0.40	24,290
2006	14,283	33,329	31,070	11,647	0.42	25,927
2007	14,927	33,483	31,213	11,843	0.44	25,085
2008	15,471	33,671	31,296	11,981	0.45	25,172
2009	15,907	34,489	32,008	12,247	0.46	25,460
2010	16,176	34,876	32,428	12,496	0.46	25,260
2011	16,370	34,678	32,073	12,559	0.47	25,139

[a] Average earnings include observations with earnings equal to zero. Wage rates are only observed for workers.
[b] Total number of observations, including observations with earnings equal to zero.

in average yearly earnings with a huge drop in earnings around the age of 60. The decline in average earnings among men may be explained by several phenomena: 1) early retirement, 2) drops in hours worked (partial retirement), 3) older people receive lower wages and 4) cohort effects. Profound cohort differences are observed among women, because of the increased female labor force participation in the last decades. We observe that a 25 year-old female earns about 17,000 euros per year on average. Around the age of 35 (when most women raise their children) earnings are relatively low, probably because of a drop in the labor force participation and/or the number of hours work. Thereafter, earnings increase and as from the age of 50 earnings decrease again.

Unemployment and part-time employment shape the earnings profile as shown in figure 4.1.[7] Figure 4.2 therefore shows the percentage of men in full-time and part-time employment over the life-cycle for different cohorts. About 70% of all men in all cohorts seem to work full-time until the age 55.[8] However, between 2001 and 2011 it seems in all cohorts about 10% of the men moved from a full-time to a part-time job. Most men seem to end up in unemployment at older ages defined as everyone not in paid-employment. About 30% is unemployed at the age of 55 and this increases to about 90% at the age of 64 for the oldest cohort. As expected, younger cohorts of men retire later.

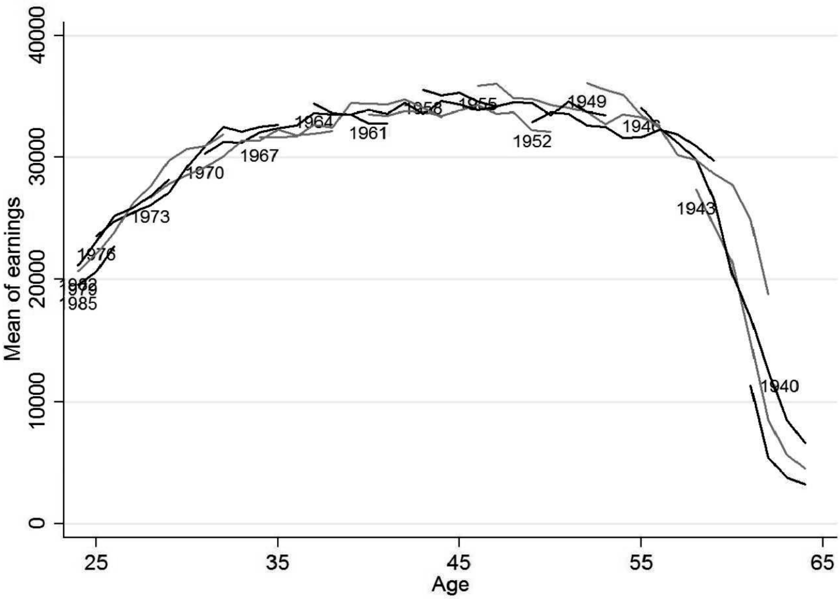
Figure 4.3 presents the percentage of women in full-time and part-time employment over the life-cycle for different cohorts. The figure indicates a substantial drop in full-time employment around the age at which women raise children. Before the age of 30 about 30-40% of women work full-time and this drops to less than 15% at the age of 40, after which it stays constant which is in line with the findings of Bosch et al. (2010). Part-time jobs, on the other hand, increase between the age of 30 and 40 from about

[7] In this paper we define people to be unemployed when they do not earn labor income from paid employment.

[8] We assume persons to be working full-time if the part-time employment factor is equal to one. Every person with a part-time employment factor of smaller than one is considered to be working part-time or unemployed. The effect of considering people with a part-time employment factor of 0.9 or bigger would be marginal as only 5.3% of men and 3.7% of women have a part-time employment factor of larger than 0.9 but smaller than one.

Figure 4.1: Life-cycle earnings of men (a) and women (b)

(a) Earnings men



(b) Earnings women

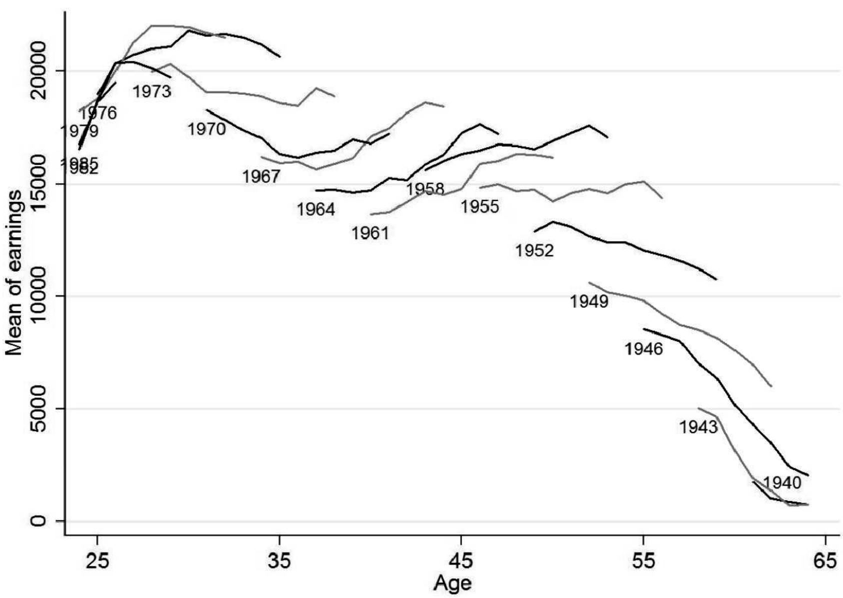
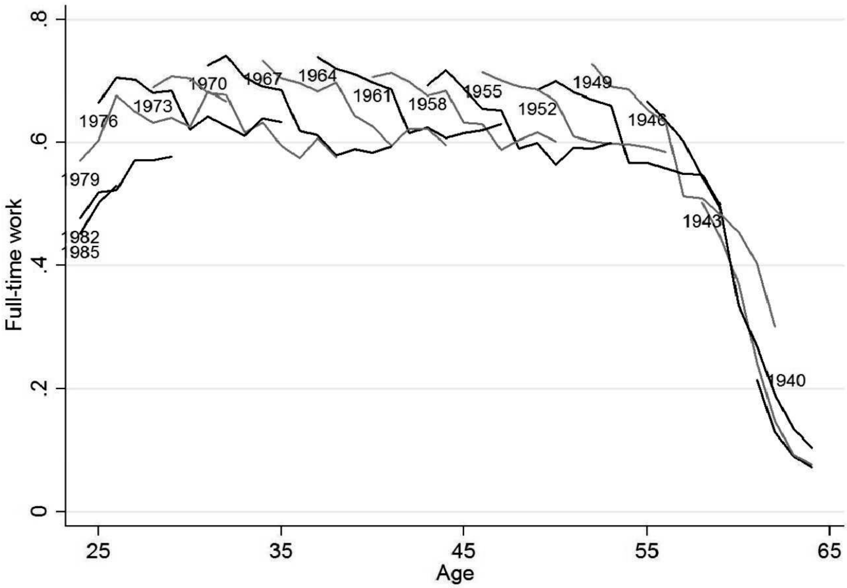
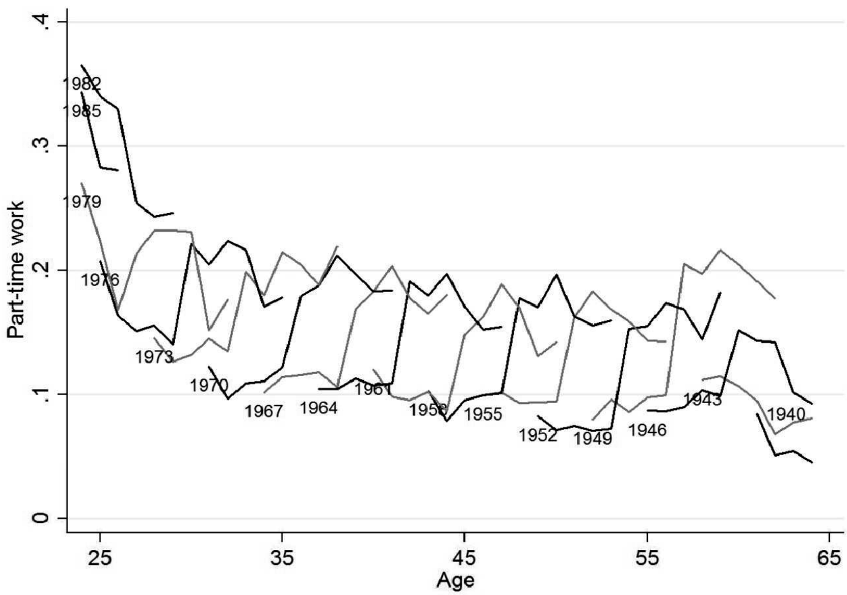


Figure 4.2: Percentage of men in full-time employment (a) and part-time employment (b)

(a) Full-time employment (%)



(b) Part-time employment (%)



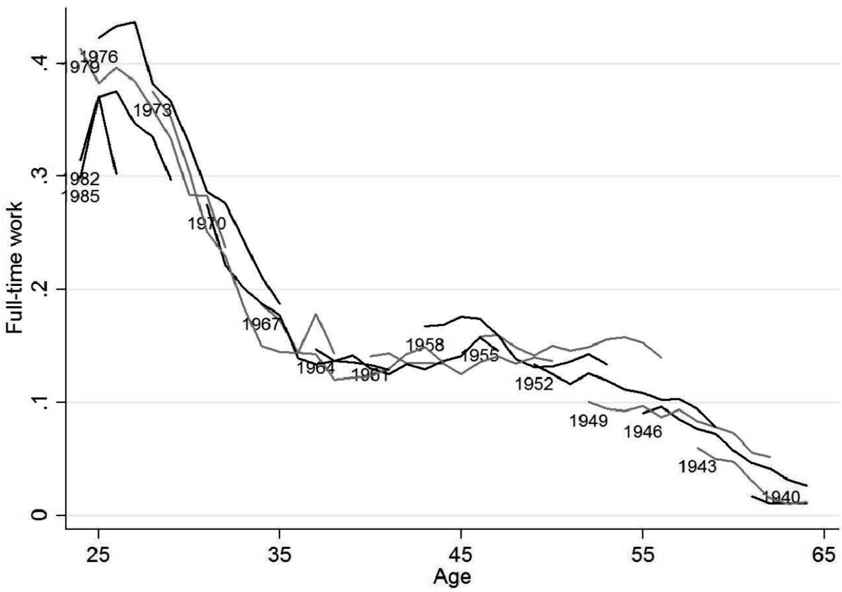
7 to 15%. Unemployment is much lower for younger cohorts than for older cohorts of women. Part-time jobs, however, increase for younger cohorts.

Figure 4.4 shows the average full-time equivalent yearly wage for those in full-time and part-time employment. Average yearly wages of men are approximately 30,000 euros at the age of 25 and about 50,000 euros at the age of 58. Female yearly average wages increase from 27,000 euros at the age of 25 to 35,000 euros at the age of 35 after which it remains relatively constant.

Decomposing the observed wages for persons in full-time and part-time employment shows that full-time wages are generally higher than the part-time wages. This applies to both men (figure 4.5) and women (figure 4.6). This observation may be explained by self-selection effects into full-time and part-time employment, e.g. persons with beneficial (observed and unobserved) characteristics tend to choose for full-time employment. The difference in full-time and part-time wages may also be well explained by the existence of a part-time wage penalty. To test the existence of selection and a part-time wage penalty, we use the model explained in the following sections.

Figure 4.3: Percentage of women in full-time employment (a) and part-time employment (b)

(a) Full-time employment (%)



(b) Part-time employment (%)

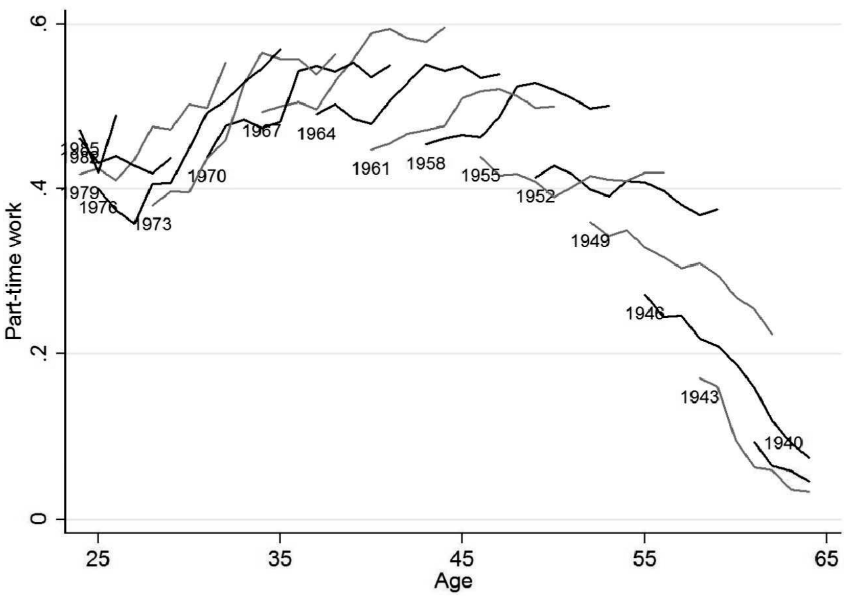
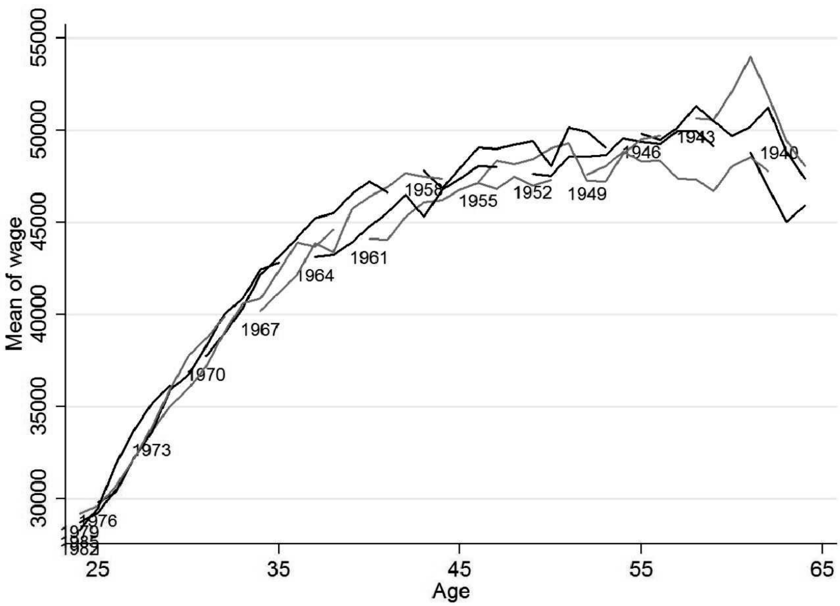


Figure 4.4: Life-cycle wages of men (a) and women (b)

(a) Wages men



(b) Wages women

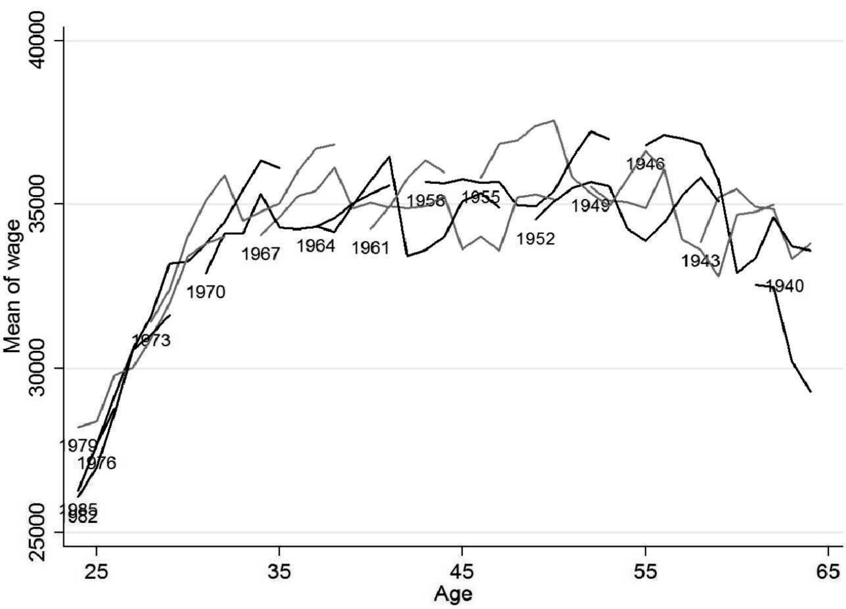


Figure 4.5: Full-time and part-time wages of men

(a) Full-time wages

(b) Part-time wages

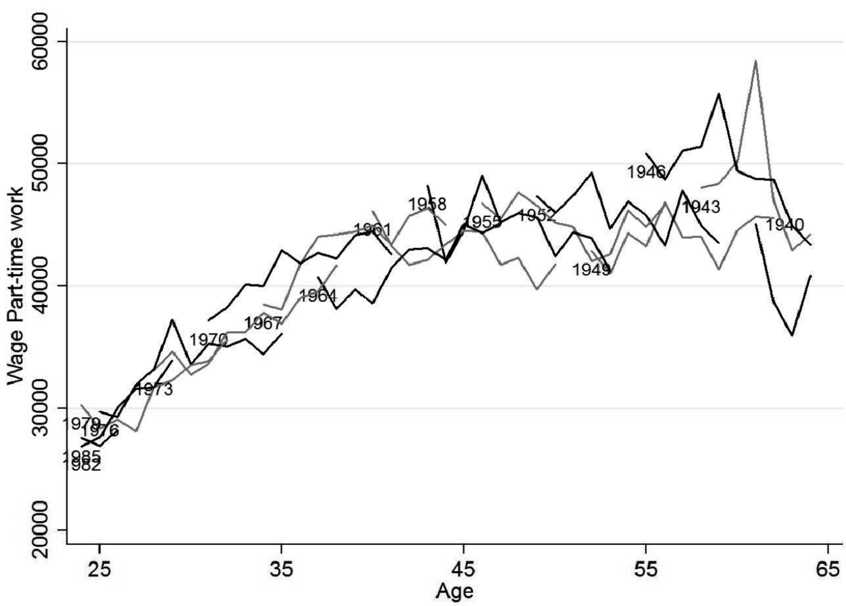
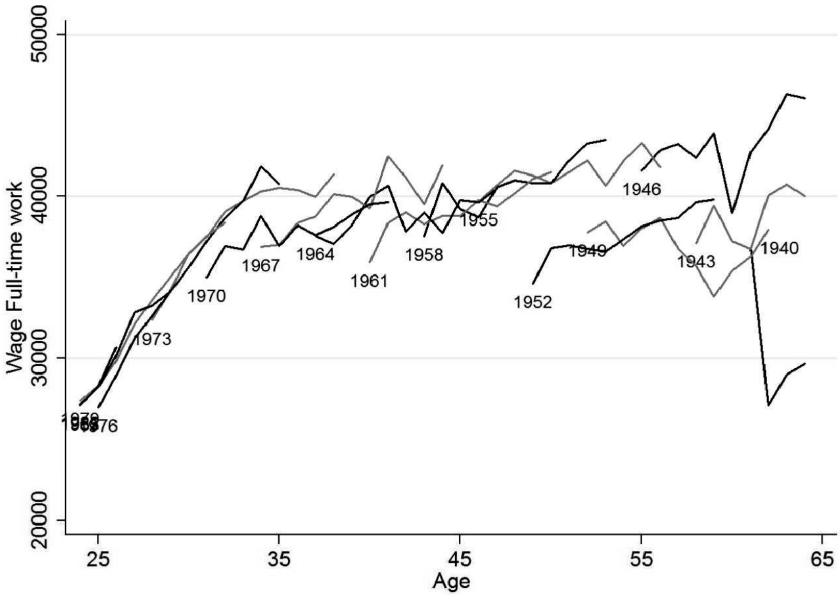
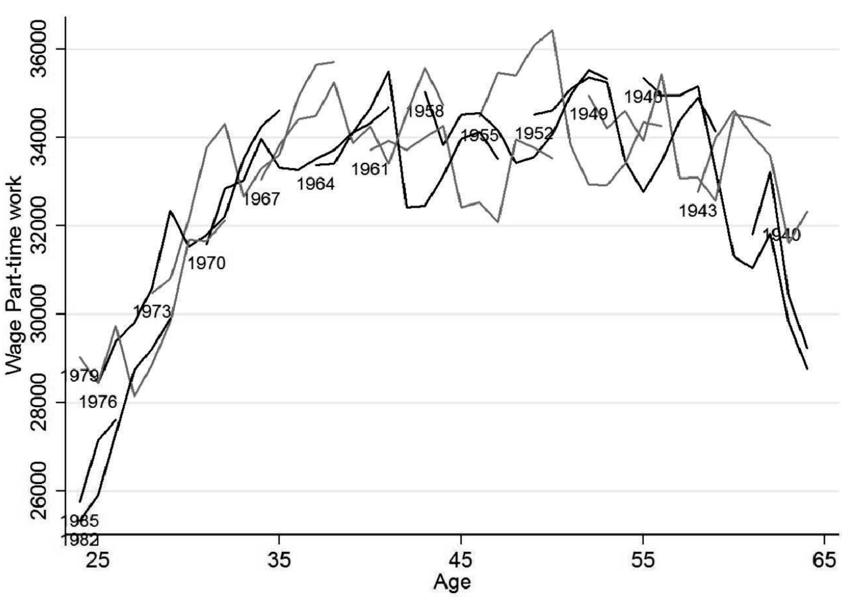


Figure 4.6: Full-time and part-time wages of women

(a) Full-time wages



(b) Part-time wages



4.3 Model

4.3.1 Panel data sample selection model

This section outlines our empirical model for analyzing wages. As discussed in section 4.2 we observe wages and the number of hours worked per year. We use a panel data sample selection model to model both wages and labor force participation at the extensive and intensive margin. The model can be written as follows:

$$y_{it}^* = x_{it}\beta + \alpha_i + u_{it} \quad i = 1, ..., N \quad t = 1, ..., T \tag{4.1}$$

$$h_{it}^* = z_{it}\gamma + \eta_i + v_{it} \tag{4.2}$$

$$y_{it} = \begin{cases} y_{it}^* & \text{if } h_{it}^* > \delta_{1t} \\ \text{unobserved} & \text{otherwise} \end{cases} \tag{4.3}$$

$$h_{it} = \begin{cases} 0 \quad \text{(no participation)} & \text{if } h_{it}^* \leq \delta_{1t} \\ 1 \quad \text{(part-time)} & \text{if } \delta_{1t} < h_{it}^* \leq \delta_{2t} \\ 2 \quad \text{(part-time)} & \text{if } \delta_{2t} < h_{it}^* \leq \delta_{3t} \\ \vdots & \\ J \quad \text{(full-time)} & \text{if } \delta_{Jt} < h_{it}^* \end{cases} \tag{4.4}$$

where y_{it} is the log full-time equivalent wage for individual i in period t. h_{it} contains J categories of labor (no labor force participation, several categories of part-time labor force participation, and full-time labor force participation). Furthermore, x_{it} and z_{it} are vectors of explanatory variables. For identification z_{it} includes variables that do not appear in x_{it} (exclusion restrictions) such as information regarding marital status, children and other household characteristics. β and γ are unknown parameter vectors to be estimated and α_i and η_i are unobserved individual specific effects, which are possibly correlated with x_{it} and z_{it}. Finally, δ_{jt} with $j = \{1, .., J\}$ are cut-off points to be estimated and u_{it} and v_{it} are unobserved disturbances, presumably not independent of each other,[9] which are assumed

[9] If u_{it} and v_{it} are independent, we do not need to worry about selection effects in the wage equation.

to follow a normal distribution with mean zero and variances $\sigma_{u,t}$ and $\sigma_{v,t}$. u_{it} and v_{it} are assumed to be uncorrelated with x_{it} and z_{it}.

To estimate the model we built upon the approaches of Rochina-Barrachina (1999) and Kalwij (2003). Kalwij (2003) proposed a new estimator for a panel data Tobit model in which the unobserved individual specific effects are allowed to correlate with the explanatory variables. The paper of Rochina-Barrachina (1999) is concerned with the estimation of a panel data sample selection model where both the selection and the regression equation contain individual effects allowed to be correlated with the observable variables.

Following Mundlak (1978) we parameterize the individual specific effect in the selection equation (4.2) as a linear function of the average explanatory variables over time plus a random individual specific effect that is assumed to be independent of the explanatory variables:[10]

$$\eta_i = \overline{z_i}\theta + c_i \tag{4.5}$$

where c_i is a random effect that is assumed to be a normally distributed random variable with mean zero and variance σ_c. Substituting (4.5) into (4.2) yields:

$$h^*_{it} = z_{it}\gamma + \overline{z_i}\theta + \mu_{it} \tag{4.6}$$

where $\mu_{it} = c_i + v_{it}$. Given the distributional assumptions it holds that $\mu_{it} \sim N(0, \sigma_{\mu,t})$, where $\sigma^2_{\mu,t} = \sigma^2_c + \sigma^2_{v,t}$. Furthermore, μ_{it} is allowed to be serially dependent (this is important, because of the term c_i).

By taking first- and higher order differences we eliminate the individual specific unobserved effects α_i without having to assume a specific parameterization of the individual unobserved effect in the wage equation (4.1). We can only observe wage differences for those observations for

[10] An application of Mundlak (1978) to panel data selection models was first used in Wooldridge (1995).

which an individual has worked at both time t and $t-m$:

$$y_{it} - y_{it-m} = \begin{cases} y^*_{it} - y^*_{it-m} & \text{if } h^*_{it-m} > \delta_{1t-m} \text{ and } h^*_{it} > \delta_{1t} \\ \text{unobserved} & \text{otherwise} \end{cases} \tag{4.7}$$

where

$$y^*_{it} - y^*_{it-m} = (x_{it} - x_{it-m})\beta + (u_{it} - u_{it-m}), m \geq 1 \tag{4.8}$$

Estimating equation (4.8) by OLS would yield inconsistent estimates of β as the conditional expectation of the error term is unlikely to be zero due to correlation between u_{it} and v_{it} (e.g. selection effects into work). Therefore, Rochina-Barrachina (1999) calculates the expectation conditional on $h^*_{it-m} > \delta_{1t-m}$ and $h^*_{it} > \delta_{1t}$.[11]

$$\begin{aligned} & E[y_{it} - y_{it-m} | x_i, z_i, h^*_{it-m} > \delta_{1t-m}, h^*_{it} > \delta_{1t}] \\ & = (x_{it} - x_{it-m})\beta + E[u_{it} - u_{it-m} | x_i, z_i, h^*_{it-m} > \delta_{1t-m}, h^*_{it} > \delta_{1t}] \\ & = (x_{it} - x_{it-m})\beta \\ & + E[u_{it} - u_{it-m} | x_i, z_i, \mu_{it-m} > \delta_{1t-m} - z_{it-m}\gamma - \overline{z_i}\theta, \mu_{it} > \delta_{1t} - z_{it}\gamma - \overline{z_i}\theta] \end{aligned} \tag{4.9}$$

The errors $[(u_{it} - u_{it-m}), \mu_{it-m}, \mu_{it}]$ are assumed to be trivariate normally distributed conditional on x_i and z_i. Denote the correlation coefficient of μ_{it-m} and μ_{it} by ρ_{tm}. By taking the derivative of the moment generating function of the truncated multi-normal distribution with respect to t and evaluating the function in $t = 0$, Rochina-Barrachina (1999) obtains the following conditional mean of the error term $(u_{it} - u_{it-m})$:[12]

$$\begin{aligned} & E[u_{it} - u_{it-m} | \mu_{it-m} > \delta_{1t-m} - z_{it-m}\gamma - \overline{z_i}\theta, \mu_{it} > \delta_{1t} - z_{it}\gamma - \overline{z_i}\theta] \\ & = \pi_{1tm}\lambda_{1itm}(M_{it-m}, M_{it}, \rho_{tm}) + \pi_{2tm}\lambda_{2itm}(M_{it-m}, M_{it}, \rho_{tm}) \end{aligned} \tag{4.10}$$

[11] The method of Rochina-Barrachina (1999) is a specific case of our general model presented in equations (4.1)-(4.4) in which only information on work versus no work is used. Equation (4.4) contains two categories: no participation and participation.

[12] This result is based on calculating the first moment of the truncated multivariate normal distribution as in Tallis (1961).

where

$$M_{it-m} = (-\delta_{1t-m} + z_{it-m}\gamma + \overline{z_i}\theta)/\sigma_{\mu,t-m} \quad (4.11)$$

$$M_{it} = (-\delta_{1t} + z_{it}\gamma + \overline{z_i}\theta)/\sigma_{\mu,t} \quad (4.12)$$

and

$$\lambda_{1tm}(M_{it-m}, M_{it}, \rho_{tm}) = \frac{\phi(M_{it-m})\Phi\left((M_{it} - \rho_{tm}M_{it-m})/\sqrt{1-\rho_{tm}^2}\right)}{\Phi^2(M_{it-m}, M_{it}; \rho_{tm})} \quad (4.13)$$

$$\lambda_{2tm}(M_{it-m}, M_{it}, \rho_{tm}) = \frac{\phi(M_{it})\Phi\left((M_{it-m} - \rho_{tm}M_{it})/\sqrt{1-\rho_{tm}^2}\right)}{\Phi^2(M_{it-m}, M_{it}; \rho_{tm})} \quad (4.14)$$

Applying OLS on the sample of first- and higher order differences will yield consistent estimates of β if the selection correction terms (4.10) are added to (4.1). If added to the regression equation, the new error term $\xi_{it} \equiv (u_{it} - u_{it-m}) - (\pi_{1tm}\lambda_{1itm} + \pi_{2tm}\lambda_{2itm})$ has a conditional expectation of zero by construction.

Panel data sample selection model with part-time employment 4.3.2

The proposed method by Rochina-Barrachina (1999) takes into account the binary selection of work versus no work. We argue that more information regarding the correlation between u_{it} and v_{it} can be added to the model by additionally taking into account labor supply at the intensive margin.

By using an ordered selection equation instead of a binary selection equation, we are able to take into account the extra information available from observing part-time and full-time work. Thus, instead of only correcting for systematic differences between those who work and those who do not work, we also take into account unobserved differences between those who work part-time and full-time.

We extend equation (4.9) by taking into account the lower- and upper thresholds of working hours categories, which yields

$$\begin{aligned}
&E[y_{it} - y_{it-m}|x_i, z_i, \delta_{j,t} < h^*_{it} \leq \delta_{j+1,t}, \delta_{j,t-m} < h^*_{it-m} \leq \delta_{j+1,t-m}] \\
&= (x_{it} - x_{it-m})\beta \\
&+ E[u_{it} - u_{it-m}|x_i, z_i, \delta_{j,t} < h^*_{it} \leq \delta_{j+1,t}, \delta_{j,t-m} < h^*_{it-m} \leq \delta_{j+1,t-m}] \\
&= (x_{it} - x_{it-m})\beta \\
&+ E[u_{it} - u_{it-m}|x_i, z_i, G_{it-m} \leq \mu_{it-m} < H_{it-m}, G_{it} \leq \mu_{it} < H_{it}] \quad (4.15)
\end{aligned}$$

where

$$H_{it-m} = -\delta_{j,t-m} + z_{it-m}\gamma + \overline{z_i}\theta \quad (4.16)$$

$$G_{it-m} = -\delta_{j+1,t-m} + z_{it-m}\gamma + \overline{z_i}\theta \quad (4.17)$$

$$H_{it} = -\delta_{j,t} + z_{it}\gamma + \overline{z_i}\theta \quad (4.18)$$

$$G_{it} = -\delta_{j+1,t} + z_{it}\gamma + \overline{z_i}\theta \quad (4.19)$$

and where j is the working hours category of individual i at time t. For persons who do not work at time t, $\delta_{0,t} = -\infty$. For these people, $H_{it} = \infty$. Similarly, for persons engaged in full-time work at time t, $\delta_{J+1,t} = \infty$ such that $G_{it} = -\infty$.

As in the framework of Rochina-Barrachina (1999), the errors $[(u_{it} - u_{it-m}), \mu_{it-m}, \mu_{it}]$ are assumed to be trivariate normally distributed conditional on x_i and z_i. Denote the correlation coefficient of μ_{it-m} and μ_{it} by ρ_{tm}. We can write out the conditional mean in (4.15) by:

$$\begin{aligned}
&\mathrm{E}(u_{it} - u_{it-m}|x_i, z_i, G_{it-m} \leq \mu_{it-m} < H_{it-m}, G_{it} \leq \mu_{it} < H_{it}) = \quad (4.20) \\
&\pi_{1tm}\lambda_{1itm}(\rho_{tm}, b_{it}, a_{it-m}, b_{it-m}) \\
&+ \pi_{2tm}\lambda_{2itm}(\rho_{tm}, a_{it}, a_{it-m}, b_{it-m}) \\
&+ \pi_{3tm}\lambda_{3itm}(\rho_{tm}, a_{it}, b_{it}, b_{it-m}) \\
&+ \pi_{4tm}\lambda_{4itm}(\rho_{tm}, a_{it}, b_{it}, a_{it-m})
\end{aligned}$$

where

$$a_{it-m} = \frac{G_{it-m}}{\sigma_{\mu,t-m}} \tag{4.21}$$

$$b_{it-m} = \frac{H_{it-m}}{\sigma_{\mu,t-m}} \tag{4.22}$$

$$a_{it} = \frac{G_{it}}{\sigma_{\mu,t}} \tag{4.23}$$

$$b_{it} = \frac{H_{it}}{\sigma_{\mu,t}} \tag{4.24}$$

with $\sigma_{\mu,t}$ and with $\sigma_{\mu,t-m}$ being the variances of the error term of the selection equation for time t and $t-m$ respectively and where

$$\lambda_{1itm}(\rho_{tm}, b_{i,t}, a_{it-m}, b_{it-m}) = \frac{\phi(b_{it})\left[\Phi\left((b_{it-m} - \rho_{tm} b_{it}) / \sqrt{1-\rho_{tm}^2}\right) - \Phi\left((a_{it-m} - \rho_{tm} b_{it}) / \sqrt{1-\rho_{tm}^2}\right)\right]}{\Phi^2(b_{it-m}, b_{it}; \rho_{tm}) - \Phi^2(a_{it-m}, a_{it}; \rho_{tm})} \tag{4.25}$$

$$\lambda_{2itm}(\rho_{tm}, a_{it}, a_{it-m}, b_{it-m}) = \frac{\phi(a_{it})\left[\Phi\left((b_{it-m} - \rho_{tm} a_{it}) / \sqrt{1-\rho_{tm}^2}\right) - \Phi\left((a_{it-m} - \rho_{tm} a_{it}) / \sqrt{1-\rho_{tm}^2}\right)\right]}{\Phi^2(b_{it-m}, b_{it}; \rho_{tm}) - \Phi^2(a_{it-m}, a_{it}; \rho_{tm})} \tag{4.26}$$

$$\lambda_{3itm}(\rho_{tm}, a_{it}, b_{it}, b_{it-m}) = \frac{\phi(b_{it-m})\left[\Phi\left((b_{it} - \rho_{tm} b_{it-m}) / \sqrt{1-\rho_{tm}^2}\right) - \Phi\left((a_{it} - \rho_{tm} b_{it-m}) / \sqrt{1-\rho_{tm}^2}\right)\right]}{\Phi^2(b_{it-m}, b_{it}; \rho_{tm}) - \Phi^2(a_{it-m}, a_{it}; \rho_{tm})} \tag{4.27}$$

$$\lambda_{4itm}(\rho_{tm}, a_{it}, b_{it}, a_{it-m}) = \frac{\phi(a_{it-m})\left[\Phi\left((b_{it} - \rho_{tm} a_{it-m}) / \sqrt{1-\rho_{tm}^2}\right) - \Phi\left((a_{it} - \rho_{tm} a_{it-m}) / \sqrt{1-\rho_{tm}^2}\right)\right]}{\Phi^2(b_{it-m}, b_{it}; \rho_{tm}) - \Phi^2(a_{it-m}, a_{it}; \rho_{tm})} \tag{4.28}$$

For the derivation of this result by calculating the first moment of the doubly truncated multivariate distribution, we refer to Appendix 4.A. As

$\xi_{itm} \equiv (u_{it} - u_{it-m}) - (\pi_{1tm}\lambda_{1itm} + \pi_{2tm}\lambda_{2itm} + \pi_{3tm}\lambda_{3itm} + \pi_{4tm}\lambda_{4itm})$ has a conditional expectation of zero by construction, taking into account both the lower- and upper thresholds of working hours categories results in four selection correction terms; two more than the binary selection approach of Rochina-Barrachina (1999).[13]

4.3.3 Experience and unemployment

Labor market experience has a positive return on the wage rate (see for example Dustmann and Meghir 2005). On the other hand, unemployment has a negative effect on post-unemployment wages (see for example Schmieder et al. 2013). In our proposed model in section 4.3.2, we are able to take into account information regarding experience by investigating wage differences between t and $t - m$ ($m = \{1, 2, 3, ..., 10\}$). When people experience years of unemployment between t and $t - m$ we take this explicitly into account in the model by including a variable indicating the number of years without labor income between time t and $t - m$. This provides information about how wage growth is influenced by years of unemployment. Our large data set allows us to investigate the effect of unemployment on wage growth for men and women at different ages and during different stages of the business cycle. Since the effect of the number of years unemployed on wage growth may be nonlinear, we include the number of years unemployed as a linear spline with knots at 0, 1 and 3 years of unemployment. This linear spline takes into account that the effect of unemployment on wage growth may be different in the first year compared to the second and third year and four or more years.

4.3.4 Estimation

To estimate the model we use a two-step estimation procedure. In the first step we deal with the selection equation. We estimate the following bivariate ordered probit model for each $s = \{t, t - m\}$.

[13]Technically, this is a consequence of the difference in analyzing the first moment of a singly and doubly truncated multivariate normal distribution.

$$h^*_{it-m} = z_{it-m}\gamma_{t-m} + \overline{z_i}\theta_{t-m} + \mu_{it-m} \tag{4.29}$$

$$h^*_{it} = z_{it}\gamma_t + \overline{z_i}\theta_t + \mu_{it} \tag{4.30}$$

$$h_{is} = \begin{cases} 0 \quad \text{(no participation)} & \text{if } h^*_{is} \leq \delta_{1s} \\ 1 \quad \text{(part-time)} & \text{if } \delta_{1s} < h^*_{is} \leq \delta_{2s} \\ 2 \quad \text{(part-time)} & \text{if } \delta_{2s} < h^*_{is} \leq \delta_{3s} \\ \vdots & \\ J \quad \text{(full-time)} & \text{if } \delta_{Js} < h^*_{is} \end{cases} \tag{4.31}$$

The bivariate ordered probit model takes into account the correlation between μ_{it} and μ_{it-m}. This is necessary because we assume that this error-term has a time-constant individual component (c_i in $\mu_{it} = c_i + v_{it}$, see section 4.3).

In the second step we construct the correction terms λ_{1itm}, λ_{2itm}, λ_{3itm} and λ_{4itm} by using the estimates $\hat{a}_{it}$, $\hat{a}_{it-m}$, $\hat{b}_{it}$, $\hat{b}_{it-m}$, $\hat{\sigma}_{\mu,t}$, $\hat{\sigma}_{\mu,t-m}$ and $\hat{\rho}_{tm}$. $\hat{\lambda}_{1itm}$, $\hat{\lambda}_{2itm}$, $\hat{\lambda}_{3itm}$ and $\hat{\lambda}_{4itm}$ are used as additional regressors in the wage equation to obtain consistent estimates of β by OLS on the sample of wages observed in t and $t-m$.[14,15] $\binom{M}{2} \times 4$ selection terms are added to the wage equation.[16] We estimate

$$y^*_{it} - y^*_{it-m} = (x_{it} - x_{it-m})\beta + \sum_{c=1}^{4} \pi_{ctm}\lambda_{citm} + (u_{it} - u_{it-m}),$$
$$m \geq 1 \tag{4.32}$$

[14]We use bootstrapped standard errors for inference in the two-stage approach (Wooldridge 2002).

[15]Note that our estimation approach is slightly different from the approach taken in Rochina-Barrachina (1999). Rochina-Barrachina (1999) estimates separate OLS regressions for each s and uses a minimum distance estimator on the separate OLS regressions to obtain the regression results. We, on the other hand, estimate one OLS regression on first- and higher order differences. Both approaches assume that the effects are the same for each s.

[16]$M = 10$. The bivariate model consists of pairs of 2. We obtained 4 selection correction terms per combination.

4.4 Estimation results

4.4.1 Selection equation

We model the first-stage bivariate ordered probit models with four ordered categories of labor force participation: 1) no participation, 2) participation lower than or equal to 50% of the full-time working hours, 3) more than 50% but less than 100% of the full-time working hours, and 4) working full-time.[17]

We allow for a semi-parametric specification of age effects by using age-dummies as explanatory variables in vector z. Following Ermisch and Wright (1993) and Paci et al. (1995), we use information on marital status (dummies for married, divorced and widowed) and children (the number of children and age of the youngest child) as exclusion restrictions in z_{it}. Furthermore, we use a dummy variable that indicates whether an individual has a partner aged 62 or older. As an additional control variable we include a dummy for first-generation immigrants. $\overline{z_i}$ includes the individual's time-averages of the marital status dummies, the variables providing information on children and the dummy whether there is a partner aged 62 or older present in the household.

The bivariate ordered selection model is estimated for each combination of t and $t-m$.[18] The separate estimations capture period and cohort differences in labor force participation.

4.4.2 Wage equation

The main equation (4.1) contains a flexible semi-parametric specification of age- and period effects (following Kalwij and Alessie 2007). However, age, period, and cohort effects (captured in the individual effect) cannot be identified empirically because the calendar year is equal to the year of birth plus age thereby spanning up the vector space. To identify age, period, and cohort effects we follow the identification restriction proposed by

[17] We do a sensitivity check with more part-time employment categories.

[18] In our case with data from 2001-2011, this implies separate estimations for $(2002, 2001), (2003, 2002), (2003, 2001), ..., (2011, 2010), ..., (2011, 2001)$.

Deaton and Paxson (1993). This means that we assume that all remaining period effects add up to zero and are orthogonal to a linear time trend.

The estimated period effects (not reported here) are generally significant but their effects on wages are rather small compared to the age effects, e.g. most of the wage growth is a result of age and cohort effects.

Selection over the life-cycle

Figure 4.7 shows the estimated age coefficients of the wage regressions for men and women. The figure indicates the differences in the estimated age coefficients for 1) a model without selection correction (solid line), 2) a model with binary selection correction as in Rochina-Barrachina (1999) (dotted line) and 3) a model with ordered selection correction as proposed in this paper (dashed line).

First focusing on the model without selection (solid line), the results can be interpreted as follows. A 64 year old male has a 60% higher wage than a 24 year old male. A 64 year old female has an approximately 50% higher wage than a 24 year old female. Wages slightly decrease after the age of 58. The wage at age 64 is significantly lower than the wage at age 58.[19] The wage at age 64 is comparable to the wage at age 46.[20] Among women, we observe a lower wage growth from the age of 47 as the slope of the wage curve decreases.

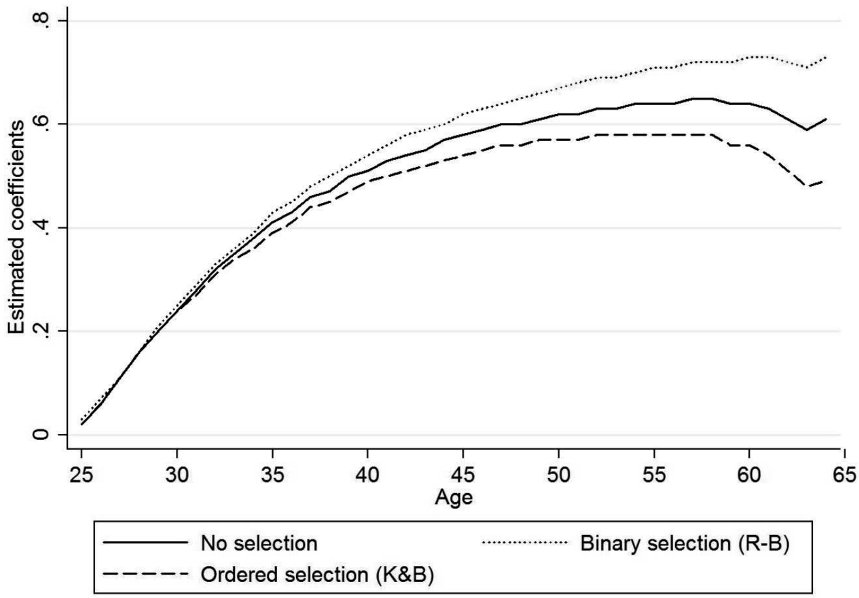
To test for selection, we follow Rochina-Barrachina (1999) who argues that a valid test of no selection is a *Wald-test* of the joint significance of the selection terms. In the binary selection model this means a Wald-test on $\binom{M}{2} \times 2$ coefficients. In the ordered selection model this means a Wald-test on $\binom{M}{2} \times 4$ coefficients.

For men, the selection correction terms of the binary selection correction are jointly significant.[21] Estimated age coefficients are higher than in the model without selection correction. This suggests that correcting the wages for persons whose wages are not observed gives a higher age effect on wages than the model without selection correction. This result suggest

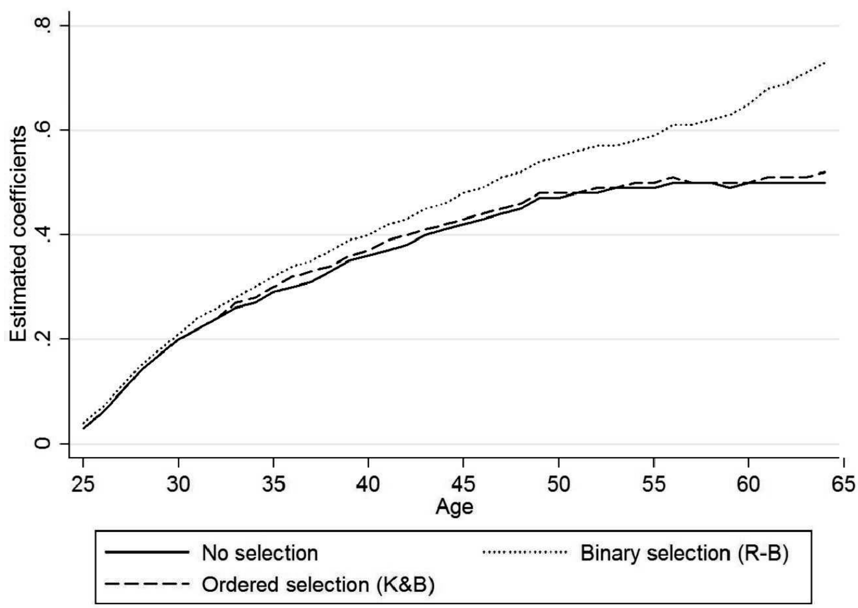
[19] $H_0 : \beta_{58} = \beta_{64}$. H_0 rejected, p-value= 0.00.
[20] $H_0 : \beta_{46} = \beta_{64}$. H_0 can not be rejected, p-value= 0.20.
[21] P-value= 0.00.

Figure 4.7: Binary versus ordered selection correction regressions and regressions without selection correction of men (a) and women (b)

(a) Estimated coefficients men



(b) Estimated coefficients women



the existence of negative selection into work over the life-cycle among men, e.g. men with worse observed and unobserved characteristics tend to work.

Testing the joint significance of the selection correction terms of the ordered selection correction are jointly significant.[22] Again, we find an inverse U-shape of wages over age that is even more pronounced than in the model with binary selection correction and the model without selection correction. The results show that, wages drop by a substantial 9%-points from the peak at age 55 to the wage at age 64. Correcting the wages for persons whose wages are not observed by the ordered selection procedure gives a lower age effect on wages than the model without selection correction. This result suggests positive selection into work over the life-cycle among men. Positive selection seems especially pronounced towards the end of the career. Such positive selection into work would have remained unnoticed in a model that corrects for selection by using a binary indicator. In stead, the model with the binary selection procedure suggests that there is negative selection into work among men. These different results indicate that correcting for selection into work and working hours simultaneously may lead to different conclusions than correcting solely for selection into work.

For women, the selection correction terms of the binary selection correction are jointly significant.[23] The estimation results show that women's wages tend to increase over the life-cycle. Furthermore, we find that correcting for selection with the binary selection indicator suggests the existence of negative selection into work.

Testing the joint significance of the selection correction terms for the ordered selection model indicates that selection is present.[24] Whereas the model with the binary selection indicator suggests the existence of negative selection among women, the model with the ordered selection rule suggests that this negative selection is much smaller. Especially among older women.

[22] P-value= 0.00.
[23] P-value= 0.00.
[24] P-value= 0.00.

Table 4.2: Effect of career breaks (years) on wage

	Men		**Women**	
	Coeff.	S.D.	Coeff.	S.D.
Career break = 1	-0.11***	0.01	-0.07***	0.01
1 < Career break ≤ 3	-0.05***	0.01	-0.05***	0.01
Career break > 3	0.01	0.02	0.02	0.02

Based on a binary selection indicator we would conclude that negative selection over the life-cycle is present among men and women. However, adding information from working hours decisions makes us conclude that positive selection exists over the life-cycle among men while the negative selection for women is much smaller than suggested by the model with binary selection terms.

Career breaks

The effects of a career break, defined as a year in which one does not receive labor income, on the life-cycle wage is estimated by a linear spline for 1, 2-3 and 4+ years of a career break. A semi-parametric linear spline is used because of possible non-linear effects, e.g. the effect a the first year may be different from the effect of 4+ years.

The estimated coefficients of the linear spline function of career breaks are shown in table 4.2 and can be interpreted as follows. Males who suffer from a career break of at most one year have a 11% lower wage than men without a career break. Men who suffer from a career break of at most 2-years (not necessarily subsequently) face an additional 5% lower wage. An additional third year lowers the wage with another 5%. An additional year after 3 years does not significantly reduce the wage anymore.

Women with a one-year career break face a 7% lower wage than women without a career break. An additional second and third year reduce the wage by 5%. The effects of a career break are not significantly different from zero thereafter.

Education-specific selection in full-time and part-time employment 4.4.3

In this section, we estimate education-specific life-cycle wage profiles. To analyze selection into full-time and part-time wages, we estimate separate wage equations for full-time (β_{FT}) and part-time work (β_{PT}).

Taking into account education may be relevant as wage growth may differ between educational levels (Connolly and Gottschalk 2006). To analyze whether selection into full-time and part-time wages differs between educational levels, we estimate separate wage equations for different educational levels. We use the international ISCED3 standard to define low education (ISCED3= 1 or ISCED3= 2) and high education (ISCED3= 3).

In the selection equation, we use a simplified version of the specification of the bivariate ordered probit model due to the loss of observations when using educational information as explained in section 4.2. Basically, we no longer use a semi-parametric specification of the age effects but assume the effects of age on working hours to be quadratic. Furthermore, we assume that the age effects on hours decisions may differ between low- and high education. The same exclusion restrictions are used as in the earlier specification of the bivariate ordered probit model.

The results (figure 4.8) suggest that there is positive selection in part-time and full-time employment among men. For low-educated men we find positive selection in both part-time[25] and full-time employment.[26] However, we do not find any wage growth over the life-cycle in the part-time wage profile for low-educated men as the estimated age-coefficients are not significantly different from zero. Among high-educated men, we only find significant positive selection into full-time employment.[27] Selection into part-time employment is not significant among high-educated men.[28]

Analyzing the selection effects into part-time and full-time employment among low-educated women shows that there is positive selection into

[25]P-value= 0.00.
[26]P-value= 0.02.
[27]P-value= 0.03.
[28]P-value= 0.11.

both part-time work[29] and full-time work[30] (figure 4.9). When correcting for selection, all age effects on wages become insignificant among low-educated women working part-time, e.g. there is no significant wage growth over the life-cycle. Corrections for selection into part-time and full-time employment among high-educated women shows positive selection in both part-time[31] and full-time work[32] (figure 4.8).

4.4.4 Education-specific part-time wage penalties

Noticeable observations in figures 4.8 and 4.9 are 1) wage growth is steeper over the life-cycle for full-time employment and 2) wage growth is steeper for high-educated persons. This is true for both men and women.

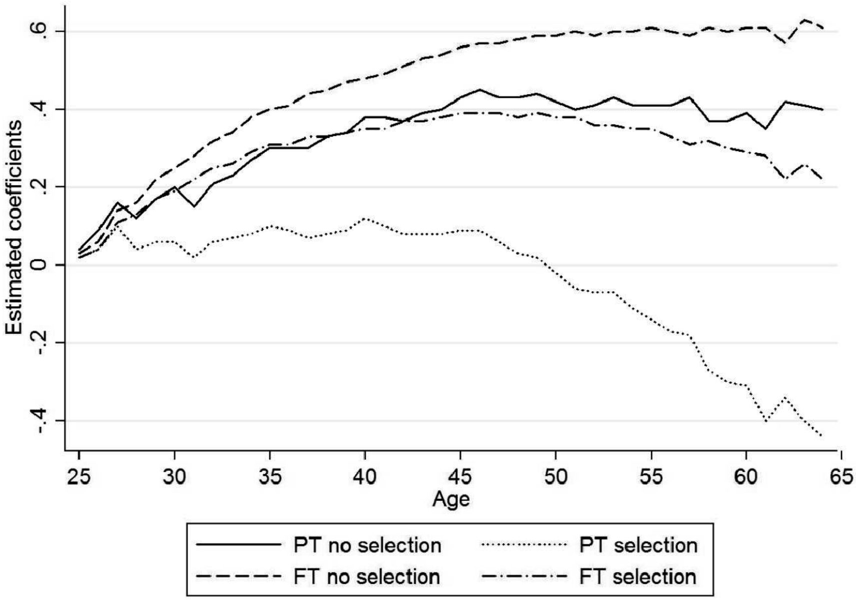
The seminal work of Mincer (1974) focused on the relationship between education and wage. Since Mincer (1974) many more advanced techniques to identify the causal link between education and wages have emerged as summarized by Card (1999) and virtually all studies find positive returns to education. More recently Connolly and Gottschalk (2006) found that also wage growth differs among educational levels which explains our second observation in figures 4.8 and 4.9.

Our first observation suggests the existence of a part-time wage penalty. Most studies find an existing part-time wage penalty although in some papers the penalty almost disappears when controlling for differences in personal- and job characteristics (for example, Manning and Petrongolo 2008). Others studies still find a part-time wage penalty after controlling for such observables with substantial cross-country variation (for example, Gornick and Jacobs 2002). A part-time wage *premium* is also found empirically (for example, Pissarides et al. 2005). However, aforementioned studies do not control for differences in unobserved characteristics such as ability, tastes and preferences for full-time work. If there are unobserved differences between persons choosing for part-time and full-time employment, such as suggested by Hakim (1997), results of aforemen-

[29] P-value= 0.02.
[30] P-value= 0.03.
[31] P-value= 0.00.
[32] P-value= 0.00.

Figure 4.8: Part-time and full-time wage regressions for (a) low-educated and (b) high-educated men

(a) Estimated coefficients low-educated



(b) Estimated coefficients high-educated

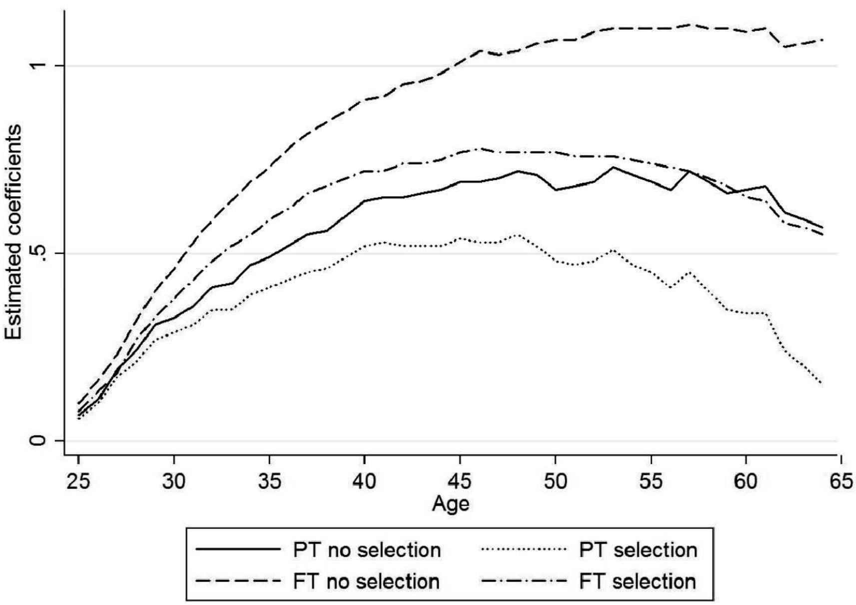
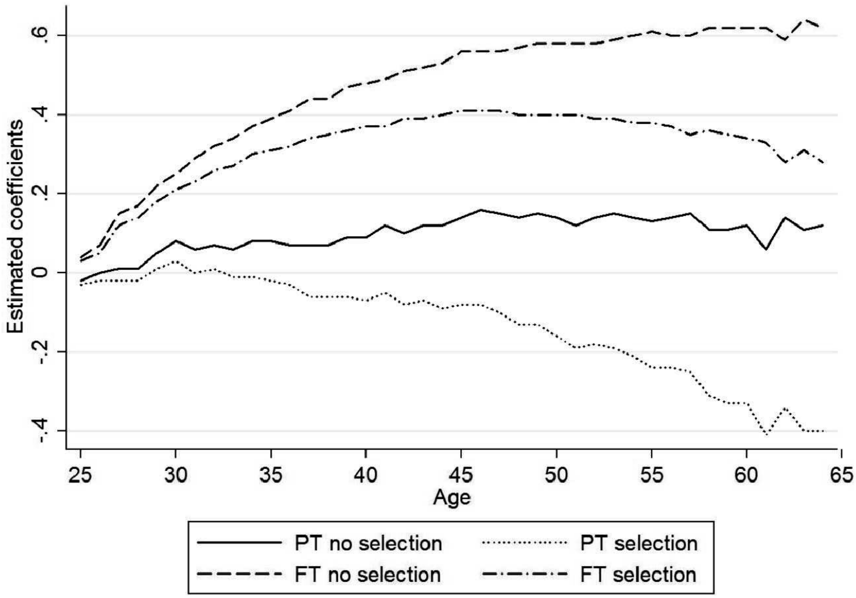
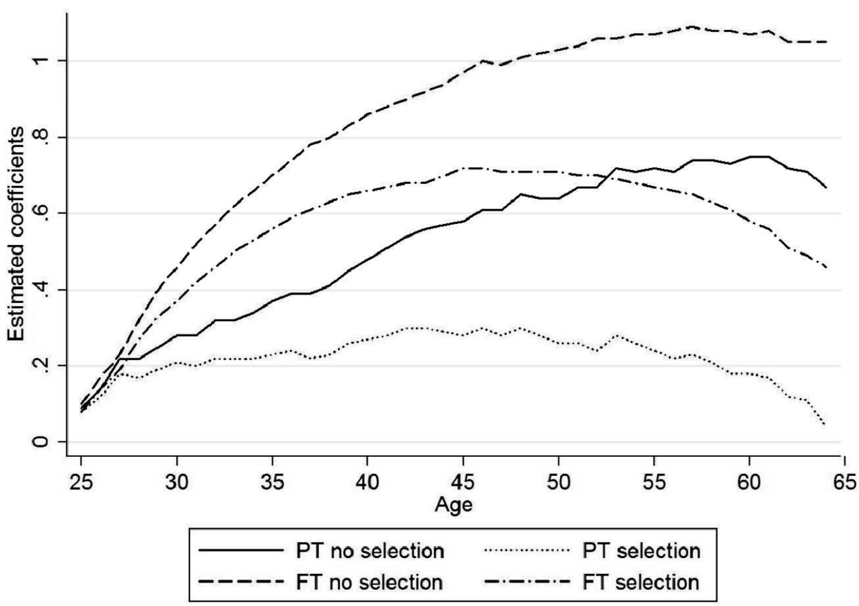


Figure 4.9: Part-time and full-time wage regressions for (a) low-educated and (b) high-educated women

(a) Estimated coefficients low-educated



(b) Estimated coefficients high-educated



tioned studies are likely to be biased as most studies on the part-time wage penalty are based on cross-sectional data and do not control for selection.[33]

Compared to these earlier studies, we take into account selection effects into part-time and full-time work. This is important as Aaronson and French (2004) and Casanova (2013) also find substantial part-time wage penalties for men (25% and 34% of a full-time wage respectively) approaching the state-pension eligible age for who part-time employment often functions as a partial retirement route. Dustmann and Schmidt (2000) do take into account selection into full-time and part-time employment when calculating wage differentials. However, their main interest lies in the wage differential between native- and migrant women and estimate these wage differentials for women in full-time and part-time employment.

Using the results in figures 4.8 and 4.9 we can test the existence of a part-time wage penalty by testing whether all differences in the age coefficients of part-time and full-time age dummies are jointly equal to zero.[34] Among high-educated men, the difference between estimated age coefficients of the part-time and full-time model is not jointly significantly different from zero meaning that we do not find a significant part-time wage penalty over the life-cycle among high-educated men.[35] We do find joint significance of the difference between estimated age coefficients for low-educated men which suggests the existence of a part-time wage penalty.[36] The part-time wage penalty is present among both low-educated[37] and high-educated[38] women.

To get an idea about the magnitude of the part-time wage penalty, we use simulation to calculate the part-time wage penalty for the mean using the following procedure: 1) we derive the unobserved heterogeneity α_i for each person, 2) we calculate the average of α_i and every variable in vector x_{it} assuming $u_{it} = 0$, 3) we set $\lambda_{1itm} = 0$, $\lambda_{2itm} = 0$, $\lambda_{3itm} = 0$ and $\lambda_{4itm} = 0$ since λ_{1itm}, λ_{2itm}, λ_{3itm} and λ_{4itm} are included in the regression to obtain

[33] Ermisch and Wright (1993) do correct for selection with cross-sectional data.
[34] $H_0 : \beta_{PT,25} - \beta_{FT,25} = ... = \beta_{PT,64} - \beta_{FT,64} = 0$.
[35] P-value= 0.45.
[36] P-value= 0.03.
[37] P-value= 0.00.
[38] P-value= 0.00.

correct estimates of β. The selection correction terms are, however, of no relevance in this simulation excercise. 4) We predict the full-time and part-time wage for the average person and 5) we calculate the differences of full-time and part-time wages for the mean and the associated variance using bootstrap.[39] The interpretation of this method is that we calculate the differences for the average person (in terms of observed and unobserved characteristics) who worked either full-time or part-time. Results are shown in table 4.3.

Table 4.3 shows the simulated part-time and full-time wages for the mean as well as the absolute difference and the relative difference (the part-time wage penalty).[40] We find that the part-time wage penalties of -16% and 14% are not statistically different from zero[41] for low-educated and high-educated men respectively (see table 4.3). Correcting for selection, career breaks and education, the size of the part-time wage penalty is negligible over the life-cycle among men. For low-educated and high-educated women we find a significant part-time wage penalty of 30% and 34% respectively (see table 4.3).

These results, ofcourse, depend on the age that is used in the simulation exercise. We use the average age that is observed in our data. Using a lower (higher) age is likely to give a smaller (larger) part-time wage penalty as the differences between full-time and part-time wages increases over the life-cycle because of cumulative effects in experience.[42] Manning and Robinson (2004), Hirsch (2005) and Russo and Hassink (2008) find that the part-time wage penalty is small or absent at the start of a career but develops over the life-cycle. This can be explained by the lower experience among part-time workers as well as by the observed lower incidence of promotions among part-time workers compared to full-time workers (Russo and Hassink 2008). These two effects tend to accumulate over

[39] Using 1500 replications.

[40] Please note that observed- and unobserved characteristics are averaged within gender and not over gender. This implies that part-time and full-time wages can be compared within gender but not between gender. Only the relative wage penalty can be compared between gender.

[41] The variance of this penalty is relatively large and therefore the difference in wage is insignificantly different from zero.

[42] In this simulation, we assume people to work either part-time of full-time during their whole working life.

Table 4.3: Part-time wage penalty for the mean for educational levels

	Men				Women			
	Low-educated		*High-educated*		*Low-educated*		*High-educated*	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
$\hat{\bar{y}}_{PT}$	47,225	12,110	48,970	7,310	29,215	4,027	36,717	3,286
$\hat{\bar{y}}_{FT}$	40,696	1,420	56,690	1,489	41,626	1,621	57,774	1,487
Absolute difference	-6,529	12,286	7,719	7,467	12,411	4,430	21,056	3,564
Relative difference	-16%		14%		30%		34%	

the life-cycle which explains the increasing gap between full-time and part-time wages (Russo and Hassink 2008). The fact that we only find a part-time wage penalty among women may be explained by compensation for the ability to combine work with care (Boeri and Van Ours 2008).

Sensitivity analyses 4.4.5

To determine the robustness of our results we perform two sensitivity analyses. Firstly, we analyze the consequences of increasing the number of part-time employment categories in the selection equation. Secondly, we discuss the possible endogeneity of the careeer breaks and the effect on the conclusions.

Increasing part-time employment categories

The paper argues that adding additional information regarding the intensive margin of participation is important in estimating wage profiles as it gives more information regarding unobserved characteristics that remain unnoted in selection correction models that only take into account the extensive margin. To prove this, we also compare our baseline results ($J = 4$) with an extended model with more part-time employment categories ($J = 7$).

We increase the number of ordered categories in the selection equation to $J = 7$: 1) no participation, 2) full-time factor between 0% and 20%, 3) full-time factor between 20% and 40%, 4) full-time factor between 40%

and 60%, 5) full-time factor between 60% and 80%, 6) full-time factor between 80% and 100%, and 7) working full-time (100%). The percentage of men observed in these categories is 26%, 0.4%, 1%, 2%, 2%, 10% and 59% respectively. The percentage of women observed in these categories is 42%, 1%, 5%, 12%, 12%, 12% and 16% respectively.

Figure 4.10 shows the estimation results when we take into account 7 working hours categories in stead of 4 in part-time and full-time wage equations.

The estimation results are highly comparable for the full-time wage profile. We observe that using $J = 7$ in stead of $J = 4$ in the first-stage causes the part-time wage profile to increase for men and decrease for women. So, the number of hours categories taken into account does matter for the second-stage wage profiles. For future research, we would like to increase J as long as there are a sufficient amount of observations in each j.

Career breaks

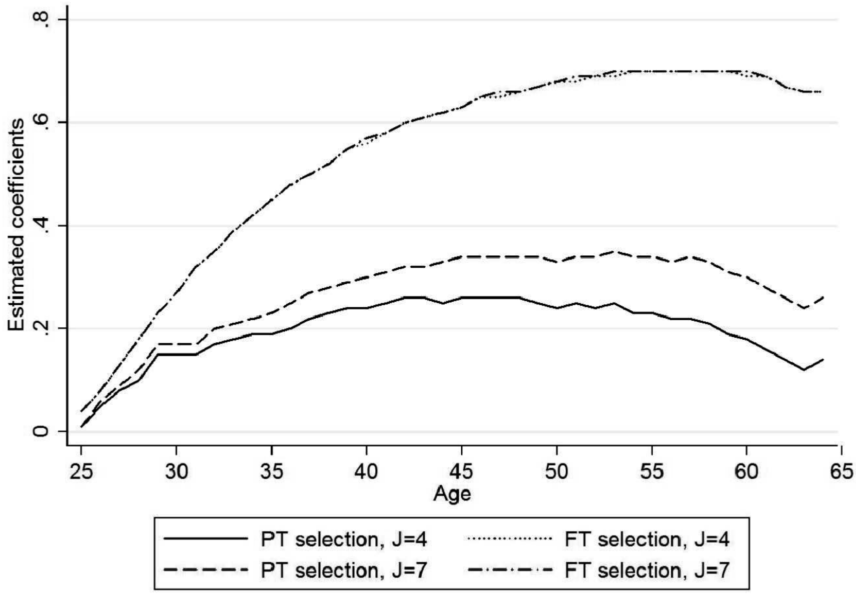
Since perons with a low-wage potential may also be subject to a career break more often (because of unemplyment, for example), the coefficients of the effects of career breaks on wages should be interpreted with caution. To check that the possible endogeity of the career breaks does not affect the main conclusions of the paper, we estimated a model excluding career break variables. Comparing the estimated wage profiles without career breaks to the model including career break variables indicates that the estimated age coefficients are highly similar. Therefore, we conclude that the inclusion of career break variables does not affect the main conclusions of the paper despite the possible endogenity of the variables.

4.5 Conclusion

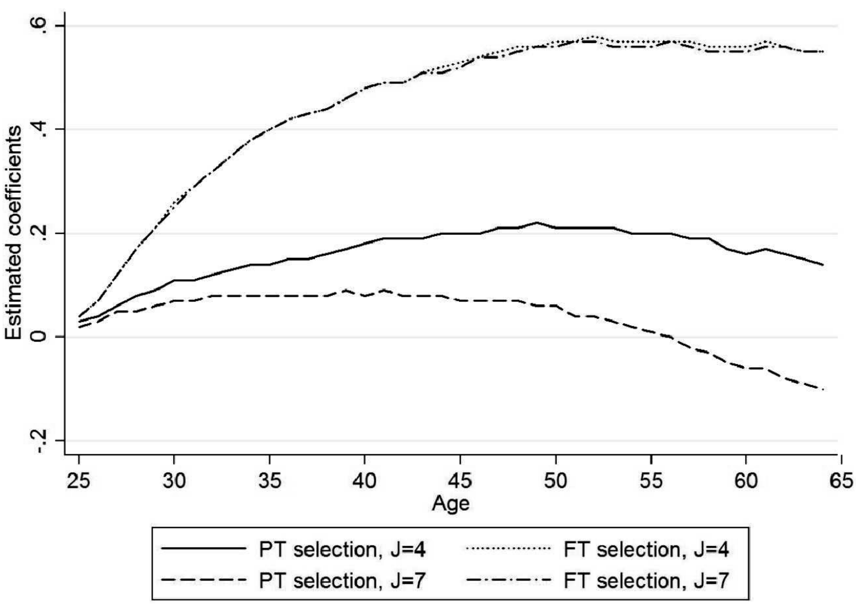
To gain insight in consumption and savings behavior over the life-cycle and to assess the adequacy of retirement savings, it is important to model life-cycle earnings as labor income usually is the primary source of income

Figure 4.10: Ordered selection correction regressions with J=7 for men (a) and women (b)

(a) Estimated coefficients men



(b) Estimated coefficients women



(Scholz et al. 2006). Also, earnings are often directly related to the accumulation of (occupational) pension rights over the life-cycle. Conclusions regarding the adequacy of retirement savings depend on a correct specification of the wage equation. However, most life-cycle models neglect the selection into wages while selection into work is likely to be nonrandom (Casanova 2013).

This paper proposes a new estimator to estimate life-cycle wage profiles using a panel data sample selection model that takes into account information about part-time and full-time work. Our proposed new estimator is an extension of the method proposed by Rochina-Barrachina (1999). Rochina-Barrachina (1999) proposes a binary selection equation to correct for selection into work. We propose an ordered selection equation to correct for selection into work and the number of hours of work simultaneously. By taking into account the number of hours that people work, extra information is available about unobserved characteristics in the wage equation. This is especially relevant for the analysis of wages over the life-cycle as women who work full-time or have a large part-time job during the upbringing of young children may have different unobserved characteristics compared to women in small part-time jobs. Also, men who retire partially may be a selective group with different observed and unobserved characteristics than men who do not retire gradually. The estimator proposed in this paper is applied to estimate life-cycle wage profiles and to analyze selection into part-time and full-time employment as well as the part-time wage penalty over the life-cycle conditional on possible career breaks.

Using the binary selection correction proposed by Rochina-Barrachina (1999) we find negative selection into work over the life-cycle among men and women. However, adding information regarding hours decisions by using the ordered selection correction proposed in this paper we find positive selection into work over the life-cycle among men and less substantial negative selection among women. This difference indicates that it is important to take into account both participation and hours decisions to account for unobserved heterogeneity in wages. This is strengthened by our analysis that increases the number of hours categories. The positive selection suggests that persons with more affluent observed and unobserved characteristics tend to work over the life-cycle whereas persons with less beneficial observed and unobserved characteristics are less likely to be employed. Career breaks have a substantial negative effect on life-cycle wages with an effect of 11% (men) and 7% (women) of the first year which increases up to 21% (men) and 17% (women) from the third year.

Education-specific life-cycle wage profiles for low- and high-educated persons show that both selection effects and part-time wage penalties may differ between these groups. Among men, we generally find positive selection. The part-time wage penalty over the life-cycle is not significantly different from zero for low- and high-educated men. Positive selection into part-time and full-time employment is found among both low-educated and high-educated women. Estimating the life-cycle wage profiles separately for low- and high-educated women substantially gives an average part-time wage penalty of 30% and 34% for low- and high-educated women respectively.

The paper shows the existence of selection into work over the life-cycle for both men and women. This has consequences for applications in which estimating life-cycle earnings processes are crucial. The extra information regarding unobserved individual heterogeneity that the proposed estimator incorporates in estimating life-cycle wages makes it well-applicable to models that depend on life-cycle earnings processes such as life-cycle models of consumption and savings (Scholz et al. 2006), earnings inequality (Cappellari 2004) and microsimulation models of future pension accumulation (Borella 2004).

4.A Derivation of the selection terms

To derive the selection correction terms based on the ordered selection equation, we need to calculate the first moment of the doubly truncated trivariate normal distribution. To calculate this first moment, we follow the approach of Manjunath and Wilhelm (2012) using the moment generating function (*m.g.f.*) of a doubly truncated multivariate normal distribution. The m.g.f. of a doubly truncated trivariate normal distribution[43] that is truncated in **a** and **b** yields (see equation 5 in Manjunath and Wilhelm 2012)

$$m(\boldsymbol{t}) = e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}} \int_{\mathbf{a}^*}^{\mathbf{b}^*} \phi_{\alpha\boldsymbol{\Sigma}}(\boldsymbol{x})d\boldsymbol{x} \tag{4.33}$$

where $\boldsymbol{x}$ is a three-dimensional normal density $\boldsymbol{x}' = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$ with location parameter $\boldsymbol{\mu} = 0$ and covariance matrix $\boldsymbol{\Sigma}$. $\phi_{\alpha\boldsymbol{\Sigma}}(\boldsymbol{x})$ is the trivariate normal distribution defined as[44]

$$\phi_{\alpha\boldsymbol{\Sigma}}(\boldsymbol{x}) = \frac{1}{\alpha(2\pi)^{3/2}|\Sigma|^{1/2}} exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\right) d\boldsymbol{x} \tag{4.34}$$

with α being the fraction of the multivariate normal distribution after truncation, $\mathbf{a}^{*\prime} = \begin{bmatrix} a_1^* & a_2^* & a_3^* \end{bmatrix}$ and $\mathbf{b}^{*\prime} = \begin{bmatrix} b_1^* & b_2^* & b_3^* \end{bmatrix}$, such

$$a_1^* = a_1 - \boldsymbol{\Sigma t} \tag{4.35}$$

$$a_2^* = a_2 - \boldsymbol{\Sigma t} \tag{4.36}$$

$$a_3^* = a_3 - \boldsymbol{\Sigma t} \tag{4.37}$$

$$b_1^* = b_1 - \boldsymbol{\Sigma t} \tag{4.38}$$

$$b_2^* = b_2 - \boldsymbol{\Sigma t} \tag{4.39}$$

[43] $[(u_{it} - u_{it-m}), \mu_{it-m}, \mu_{it}]$ are assumed to be trivariate normally distributed conditional on x_i and z_i.

[44] See Muthen (1990) for a derivation of the doubly truncated bivariate normal distribution and Tallis (1961) for a derivation of the singly truncated multivariate normal distribution.

$$b_3^* = b_3 - \boldsymbol{\Sigma t} \tag{4.40}$$

with

$$\boldsymbol{t'} = \begin{bmatrix} t_1 & t_2 & t_3 \end{bmatrix} \tag{4.41}$$

We are interested in $\mathrm{E}(x_1|a_2 \leq x_2 \leq b_2, a_3 \leq x_3 \leq b_3)$ with $a_1 = -\infty$ and $b_1 = \infty$. Therefore, we need to take the partial derivative of the m.g.f. (equation 4.33) with respect to t_1. Using the chain rule for calculating derivatives gives

$$\frac{\partial m(\boldsymbol{t})}{\partial t_1} = e^{\frac{1}{2}\boldsymbol{t'\Sigma t}} \frac{\partial \Phi_{\alpha\boldsymbol{\Sigma}}}{\partial t_1} + \Phi_{\alpha\boldsymbol{\Sigma}} \frac{\partial e^{\frac{1}{2}\boldsymbol{t'\Sigma t}}}{\partial t_1} \tag{4.42}$$

Inserting the trivariate normal distribution and applying *Leibniz's rule* for differentiation under the integral sign we get

$$\begin{aligned}
&\frac{\partial \phi_{\alpha\boldsymbol{\Sigma}}}{\partial t_1} = \frac{\partial}{\partial t_1} \int_{a_1^*}^{b_1^*} \int_{a_2^*}^{b_2^*} \int_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(x_1, x_2, x_3) dx_3 dx_2 dx_1 = \\
&- \int_{a_2^*}^{b_2^*} \int_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(b_1^*, x_2, x_3) dx_3 dx_2 + \int_{a_2^*}^{b_2^*} \int_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(a_1^*, x_2, x_3) dx_3 dx_2 \\
&- \sigma_{12} \int_{a_1^*}^{b_1^*} \int_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(x_1, b_2^*, x_3) dx_3 dx_1 + \sigma_{12} \int_{a_1^*}^{b_1^*} \int_{a_3^*}^{b_3^*} \phi_{\alpha\boldsymbol{\Sigma}}(x_1, a_2^*, x_3) dx_3 dx_1 \\
&- \sigma_{13} \int_{a_1^*}^{b_1^*} \int_{a_2^*}^{b_2^*} \phi_{\alpha\boldsymbol{\Sigma}}(x_1, x_2, b_3^*) dx_2 dx_1 + \sigma_{13} \int_{a_1^*}^{b_1^*} \int_{a_2^*}^{b_2^*} \phi_{\alpha\boldsymbol{\Sigma}}(x_1, x_2, a_3^*) dx_2 dx_1
\end{aligned} \tag{4.43}$$

and

$$\frac{\partial e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}}{\partial t_1} = e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}} \sum_{k=1}^{3} \sigma_{1k} t_k \tag{4.44}$$

Evaluating the derivative $\frac{\partial m(\boldsymbol{t})}{\partial t_1}$ at $\boldsymbol{t} = 0$ in order to compute the first moment (E(X_1)) gives $\frac{\partial e^{\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}}{\partial t_1} = 0$, $a_1^* = a_1$, $a_2^* = a_2$, $a_3^* = a_3$, $b_1^* = b_1$, $b_2^* = b_2$ and $b_3^* = b_3$ such that

$$\begin{aligned}
\alpha \mathrm{E}(X_1) = & \frac{\partial}{\partial t_1}\bigg|_{t_1=0} \int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3} \phi_{\boldsymbol{\Sigma}}(x_1,x_2,x_3)dx_3dx_2dx_1 = \\
& - \int_{a_2}^{b_2}\int_{a_3}^{b_3} \phi_{\boldsymbol{\Sigma}}(b_1,x_2,x_3)dx_3dx_2 + \int_{a_2}^{b_2}\int_{a_3}^{b_3} \phi_{\boldsymbol{\Sigma}}(a_1,x_2,x_3)dx_3dx_2 \\
& - \sigma_{12}\int_{a_1}^{b_1}\int_{a_3}^{b_3} \phi_{\boldsymbol{\Sigma}}(x_1,b_2,x_3)dx_3dx_1 + \sigma_{12}\int_{a_1}^{b_1}\int_{a_3}^{b_3} \phi_{\boldsymbol{\Sigma}}(x_1,a_2,x_3)dx_3dx_1 \\
& - \sigma_{13}\int_{a_1}^{b_1}\int_{a_2}^{b_2} \phi_{\boldsymbol{\Sigma}}(x_1,x_2,b_3)dx_2dx_1 + \sigma_{13}\int_{a_1}^{b_1}\int_{a_2}^{b_2} \phi_{\boldsymbol{\Sigma}}(x_1,x_2,a_3)dx_2dx_1
\end{aligned} \tag{4.45}$$

Since $a_1 = -\infty$ and $b_1 = \infty$, the terms $-\int_{a_2}^{b_2}\int_{a_3}^{b_3} \phi_{\alpha\boldsymbol{\Sigma}}(b_1,x_2,x_3)dx_3dx_2$ and $\int_{a_2}^{b_2}\int_{a_3}^{b_3} \phi_{\alpha\boldsymbol{\Sigma}}(a_1,x_2,x_3)dx_3dx_2$ are zero.

$$\begin{aligned}
\alpha \mathrm{E}(X_1) = & -\sigma_{12}\phi(b_2)\int_{a_1}^{b_1}\int_{a_3}^{b_3} \phi\left(\frac{x_1-\rho_{12}b_2}{\sqrt{1-\rho_{12}^2}}, \frac{x_3-\rho_{23}b_2}{\sqrt{1-\rho_{23}^2}}, \rho_{13}\right)dx_3dx_1 \\
& + \sigma_{12}\phi(a_2)\int_{a_1}^{b_1}\int_{a_3}^{b_3} \phi\left(\frac{x_1-\rho_{12}a_2}{\sqrt{1-\rho_{12}^2}}, \frac{x_3-\rho_{23}a_2}{\sqrt{1-\rho_{23}^2}}, \rho_{13}\right)dx_3dx_1 \\
& - \sigma_{13}\phi(b_3)\int_{a_1}^{b_1}\int_{a_2}^{b_2} \phi\left(\frac{x_1-\rho_{13}b_3}{\sqrt{1-\rho_{13}^2}}, \frac{x_2-\rho_{23}b_3}{\sqrt{1-\rho_{23}^2}}, \rho_{12}\right)dx_2dx_1
\end{aligned}$$

$$+\sigma_{13}\phi(a_3)\int_{a_1}^{b_1}\int_{a_2}^{b_2}\phi\left(\frac{x_1-\rho_{13}a_3}{\sqrt{1-\rho_{13}^2}},\frac{x_2-\rho_{23}a_3}{\sqrt{1-\rho_{23}^2}},\rho_{12}\right)dx_2dx_1 \tag{4.46}$$

Here, ρ_{12}, ρ_{13} and ρ_{23} are the correlation coefficients between. We can rewrite equation (4.46) (see Manjunath and Wilhelm 2012) such that

$$\begin{aligned}\alpha\mathrm{E}(X_1) = &-\sigma_{12}\phi(b_2)\left[\Phi\left(\frac{b_3-\rho_{23}b_2}{\sqrt{1-\rho_{23}^2}}\right)-\Phi\left(\frac{a_3-\rho_{23}b_2}{\sqrt{1-\rho_{23}^2}}\right)\right]\\ &+\sigma_{12}\phi(a_2)\left[\Phi\left(\frac{b_3-\rho_{23}a_2}{\sqrt{1-\rho_{23}^2}}\right)-\Phi\left(\frac{a_3-\rho_{23}a_2}{\sqrt{1-\rho_{23}^2}}\right)\right]\\ &-\sigma_{13}\phi(b_3)\left[\Phi\left(\frac{b_2-\rho_{23}b_3}{\sqrt{1-\rho_{23}^2}}\right)-\Phi\left(\frac{a_2-\rho_{23}b_3}{\sqrt{1-\rho_{23}^2}}\right)\right]\\ &+\sigma_{13}\phi(a_3)\left[\Phi\left(\frac{b_2-\rho_{23}a_3}{\sqrt{1-\rho_{23}^2}}\right)-\Phi\left(\frac{a_2-\rho_{23}a_3}{\sqrt{1-\rho_{23}^2}}\right)\right]\end{aligned} \tag{4.47}$$

such that the first moment of X_1 in the doubly truncated trivariate normal distribution becomes

$$\mathrm{E}(X_1)=\frac{\alpha\mathrm{E}(X_1)}{\alpha}=\frac{\alpha\mathrm{E}(X_1)}{\Phi_2(b_2,b_3,\rho_{23})-\Phi_2(a_2,a_3,\rho_{23})} \tag{4.48}$$

with $\Phi_2(.)$ being the bivariate normal distribution. $\Phi_2(b_2,b_3,\rho_{23})-\Phi_2(a_2,a_3,\rho_{23})$ is the fraction of the trivariate normal distribution after truncation, e.g. a normalization of the terms in equation (4.47). The four terms in equation (4.47) are the four selection correction terms where $-\sigma_{12}$, σ_{12}, $-\sigma_{13}$ and σ_{13} are the coefficients to be estimated in the wage equation ($\pi_1 tm$, $\pi_2 tm$, $\pi_3 tm$ and $\pi_4 tm$ in equation (4.20)). a_2, a_3, b_2, b_3 and ρ_{23} are to be estimated in the first-stage selection equation (subscript 2 denoted as t and 3 denoted as $t-m$ in equation (4.15)). $a_2 = H_{it}/\sigma_t$, $b_2 = G_{it}/\sigma_t$, $a_3 = H_{it-m}/\sigma_{t-m}$, $b_3 = G_{it-m}/\sigma_{t-m}$, $\rho_{23} = \rho_{tm}$ and H_{it}, H_{it-m}, G_{it} and G_{it-m} defined as in equations (4.16) to (4.19). With x_1 being the error term of the wage equation and x_2 and x_3 being the error terms of the selection equations we get

$$\begin{aligned}&\mathrm{E}(x_1|a_2 \leq x_2 \leq b_2, a_3 \leq x_3 \leq b_3) = \\ &\mathrm{E}(u_{it} - u_{it-m}|G_{it-m} \leq \mu_{it-m} \leq H_{it-m}, G_{it} \leq \mu_{it} \leq H_{it})\end{aligned} \tag{4.49}$$

4.B First-stage regression results

Since we estimate the first-stage bivariate ordered probit model for every combination of t and $t - m$ for $t = \{2002, ..., 2011\}$ and $m = \{1, .., 10\}$, we end up having 55 different models to construct the selection correction terms λ_{1itm}, λ_{2itm}, λ_{3itm}, λ_{4itm}. We report the estimation results for the combination 2002 and 2001 and the combination 2011 and 2010 in tables 4.4-4.5 for men and women respectively. Apart from the sign and significance, the reported coefficients have no direct interpretation.

Age-effects are with respect to the baseline of age 25. We estimate the selection equations for persons born no later than 1980. As a consequence, the baseline of age-effects shifts from $t = 2006$. Coefficients should be interpreted with respect to the estimated parameters δ_{1t}, δ_{2t}, δ_{3t}, δ_{1t-m}, δ_{2t-m} and δ_{3t-m} that indicate the thresholds between the $J = 4$ labor supply categories for time t and t respectively. ρ_{tm} indicates the correlation between the error terms at time t and $t - m$ in the selection equation.

The estimation results (see tables 4.4-4.5) show that the likelihood of participation, and especially full-time work, decreases with age. This is true for both men and women, although the decline over age is relatively smaller for men than for women. Especially in the earlier years. Also, the first-stage regressions suggest that life-cycle participation decisions changed over time. The decrease in the probability to participate over the life-cycle is much larger in 2002 than in 2011. This suggests that labor force participation over the life-cycle increased over time. For women, this can also be concluded from the part-time factor in table 4.1. For men, the differences in labor force participation seem to be concentrated at the end of the career.

We find that immigrant men are significantly and substantially less likely to work (full-time). Being married is positively related to the labor force attachment among men, but only in the later years. The number of children decreases the labor force participation while the effect of having a partner of age 62 or older is usually not significant. Being married or divorced increases labor force participation among men. For women, we find a significant and substantial negative association between the labor force participation, children and having a partner of age 62 or older. Furthermore, a woman is less likely to work (full-time) if she is an immigrant, married or widowed.

A final interesting result from the first-stage equations are the estimates of ρ_{tm}. We find that $\hat{\rho}_{tm}$ decreases for higher m (e.g. the correlation between the error terms of the selection equation decreases if the period between the choices is longer). $\hat{\rho}_{tm}$ is rather constant over time in both the bivariate ordered probit model and the bivariate probit model, but $\hat{\rho}_{tm}$ is generally higher in the bivariate probit models than in the bivariate ordered probit models. However, estimating the wage model with bivariate ordered probit selection correction while using $\hat{\rho}_{tm}$ from the bivariate probit model gives highly similar results as the estimates using $\hat{\rho}_{tm}$ from the bivariate ordered probit model as presented in figure 4.7 (dashed line). Hence, the difference in the estimates of the wage profiles of the model with binary selection correction (dotted line in figure 4.7) and the ordered selection correction (dashed line figure 4.7) is not a consequence of the difference in the correlation of the error terms estimated by the two approaches.

Table 4.4: Estimation results first-stage selection equation, men

	$t = 2002$		$t - m = 2001$		$t = 2011$		$t - m = 2010$	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Age 25	ref.		ref.					
Age 26	0.00	0.03	-0.09	0.07				
Age 27	-0.12*	0.07	-0.06	0.06				
Age 28	-0.11	0.07	-0.07	0.06				
Age 29	-0.10	0.07	0.13*	0.07				
Age 30	0.05	0.07	0.02	0.06			ref.	
Age 31	-0.06	0.07	-0.04	0.06	ref.		0.11	0.08
Age 32	-0.08	0.06	-0.10*	0.06	0.11	0.07	-0.16**	0.07
Age 33	-0.17***	0.06	-0.04	0.06	-0.15***	0.07	-0.05	0.07
Age 34	-0.09	0.07	-0.07	0.06	-0.05	0.07	-0.03	0.07
Age 35	-0.16**	0.07	-0.12*	0.06	0.00	0.07	-0.22***	0.07
Age 36	-0.16**	0.07	-0.15**	0.06	-0.21***	0.07	-0.12*	0.07
Age 37	-0.24***	0.07	-0.10*	0.06	-0.12*	0.07	-0.16**	0.07
Age 38	-0.19***	0.06	-0.09	0.06	-0.17**	0.07	-0.11	0.07
Age 39	-0.16**	0.06	-0.07	0.07	-0.08	0.07	-0.19***	0.07
Age 40	-0.16**	0.07	-0.17***	0.06	-0.20***	0.07	-0.27***	0.07
Age 41	-0.22***	0.07	-0.20***	0.06	-0.22***	0.07	-0.26***	0.07
Age 42	-0.28***	0.07	-0.16**	0.06	-0.22***	0.07	-0.27***	0.07
Age 43	-0.23***	0.07	-0.27***	0.07	-0.27***	0.07	-0.22***	0.07
Age 44	-0.27***	0.07	-0.24***	0.07	-0.24***	0.07	-0.31***	0.07
Age 45	-0.28***	0.07	-0.19***	0.07	-0.31***	0.07	-0.31***	0.07
Age 46	-0.24***	0.07	-0.24***	0.07	-0.31***	0.07	-0.23***	0.07
Age 47	-0.31***	0.07	-0.23***	0.07	-0.20***	0.07	-0.15***	0.07
Age 48	-0.29***	0.07	-0.28***	0.07	-0.16**	0.07	-0.29***	0.07
Age 49	-0.39***	0.07	-0.32***	0.07	-0.32***	0.07	-0.31***	0.07
Age 50	-0.38***	0.07	-0.33***	0.07	-0.32***	0.07	-0.28***	0.07
Age 51	-0.39***	0.07	-0.29***	0.07	-0.29***	0.07	-0.29***	0.07
Age 52	-0.34***	0.07	-0.22***	0.07	-0.32***	0.07	-0.38***	0.07
Age 53	-0.36***	0.07	-0.38***	0.07	-0.33***	0.07	-0.25***	0.08
Age 54	-0.47***	0.07	-0.48***	0.07	-0.27***	0.08	-0.35***	0.07
Age 55	-0.55***	0.07	-0.46***	0.06	-0.40***	0.07	-0.38***	0.07
Age 56	-0.59***	0.07	-0.67***	0.07	-0.39***	0.08	-0.51***	0.07
Age 57	-0.76***	0.08	-0.77***	0.07	-0.52***	0.07	-0.45***	0.07
Age 58	-0.89***	0.07	-0.93***	0.07	-0.51***	0.07	-0.51***	0.07
Age 59	-1.05***	0.07	-1.21***	0.08	-0.59***	0.07	-0.62***	0.08
Age 60	-1.37***	0.08	-1.48***	0.08	-0.67***	0.08	-0.72***	0.08
Age 61	-1.78***	0.08	-1.83***	0.08	-0.82***	0.08	-0.96***	0.08
Age 62	-2.11***	0.09	-2.13***	0.09	-1.09***	0.08	-1.43***	0.08
Age 63	-2.35***	0.10	-2.45***	0.10	-1.59***	0.08	-1.69***	0.08
Age 64	-2.53***	0.10			-1.83***	0.08		
Number of children	-0.02*	0.01	-0.02**	0.01	-0.02***	0.01	-0.02**	0.01
Single	ref.		ref.		ref.		ref.	
Married	-0.02	0.03	0.01	0.03	0.19***	0.04	0.19***	0.04
Divorced	-0.09*	0.05	-0.08*	0.05	0.17***	0.05	0.18***	0.06
Widowed	0.12	0.17	-0.04	0.15	0.10	0.14	0.07	0.14
Immigrant	-0.49***	0.03	-0.49***	0.03	-0.46***	0.03	-0.47***	0.03
Partner 62+	0.11*	0.06	0.02	0.07	-0.02	0.04	0.04	0.05
F-test $\overline{z_i}$	220.86***				64.67***			
δ_{1t}	-0.90***	0.04			-0.84***	0.05		
δ_{2t}	-0.86***	0.04			-0.79***	0.05		
δ_{3t}	-0.60***	0.04			-0.37***	0.05		
δ_{1t-m}	-0.84***	0.04			-0.84***	0.05		
δ_{2t-m}	-0.79***	0.04			-0.78***	0.05		
δ_{3t-m}	-0.54***	0.04			-0.41***	0.05		
ρ_{tm}	0.97***	0.02			0.97***	0.02		
Obs.	24,129				20,886			
Chi^2	3,247				2,088			

Table 4.5: Estimation results first-stage selection equation, women

	$t = 2002$		$t - m = 2001$		$t = 2011$		$t - m = 2010$	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Age 25	ref.		ref.					
Age 26	-0.02	0.05	0.12*	0.06				
Age 27	0.04	0.07	0.04	0.06				
Age 28	-0.04	0.07	-0.03	0.06				
Age 29	-0.13*	0.07	-0.07	0.06				
Age 30	-0.13*	0.06	-0.12**	0.06			1.19***	0.08
Age 31	-0.18***	0.06	-0.19***	0.06	0.10	0.06	1.14***	0.08
Age 32	-0.31***	0.06	-0.18***	0.06	0.08	0.06	1.04***	0.08
Age 33	-0.30***	0.06	-0.30***	0.06	-0.03	0.06	1.04***	0.08
Age 34	-0.38***	0.06	-0.31***	0.06	ref.		1.05***	0.08
Age 35	-0.37***	0.06	-0.35***	0.06	0.02	0.06	1.00***	0.08
Age 36	-0.44***	0.06	-0.39***	0.06	-0.06	0.06	0.99***	0.08
Age 37	-0.46***	0.06	-0.39***	0.06	-0.05	0.06	0.94***	0.08
Age 38	-0.44***	0.06	-0.33***	0.06	-0.11*	0.06	0.84***	0.08
Age 39	-0.45***	0.06	-0.37***	0.06	-0.17***	0.06	0.97***	0.08
Age 40	-0.44***	0.07	-0.45***	0.06	-0.04	0.06	0.86***	0.08
Age 41	-0.50***	0.07	-0.37***	0.06	-0.18***	0.06	0.88***	0.07
Age 42	-0.42***	0.07	-0.33***	0.06	-0.14***	0.06	0.87***	0.08
Age 43	-0.39***	0.07	-0.38***	0.06	-0.16***	0.06	0.93***	0.08
Age 44	-0.42***	0.07	-0.52***	0.06	-0.08	0.06	0.84***	0.08
Age 45	-0.58***	0.07	-0.51***	0.06	-0.16***	0.06	0.86***	0.08
Age 46	-0.56***	0.07	-0.54***	0.07	-0.20***	0.06	0.91***	0.08
Age 47	-0.61***	0.07	-0.65***	0.06	-0.13**	0.06	0.86***	0.07
Age 48	-0.66***	0.07	-0.68***	0.07	-0.13**	0.06	0.88***	0.08
Age 49	-0.73***	0.07	-0.74***	0.07	-0.17***	0.07	0.77***	0.08
Age 50	-0.80***	0.07	-0.75***	0.07	-0.26***	0.06	0.75***	0.08
Age 51	-0.81***	0.07	-0.92***	0.07	-0.27***	0.06	0.82***	0.08
Age 52	-0.94***	0.07	-0.99***	0.07	-0.22***	0.07	0.75***	0.08
Age 53	-1.01***	0.07	-1.04***	0.07	-0.28***	0.07	0.67***	0.08
Age 54	-1.12***	0.07	-1.24***	0.07	-0.40***	0.07	0.60***	0.08
Age 55	-1.26***	0.07	-1.24***	0.07	-0.42***	0.07	0.58***	0.08
Age 56	-1.30***	0.08	-1.29***	0.08	-0.49***	0.07	0.49***	0.08
Age 57	-1.34***	0.08	-1.45***	0.08	-0.55***	0.07	0.39***	0.08
Age 58	-1.55***	0.09	-1.57***	0.09	-0.63***	0.07	0.24***	0.08
Age 59	-1.65***	0.09	-1.78***	0.09	-0.80***	0.07	0.31***	0.08
Age 60	-1.87***	0.10	-1.84***	0.10	-0.71***	0.08	ref.	
Age 61	-2.07***	0.12	-2.10***	0.10	-1.03***	0.08	-0.16***	0.09
Age 62	-2.34***	0.11	-2.65***	0.14	-1.16***	0.08	-0.38***	0.09
Age 63	-2.69***	0.14	-2.87***	0.18	-1.47***	0.08	-0.67***	0.09
Age 64	-3.14***	0.22			-1.84***	0.09		
Number of children	-0.08***	0.01	-0.10***	0.01	-0.08***	0.01	-0.09***	0.01
Single	ref.		ref.		ref.		ref.	
Married	-0.23***	0.04	-0.21***	0.04	0.01	0.04	0.01	0.05
Divorced	-0.09	0.06	-0.12**	0.05	0.09	0.06	0.14**	0.06
Widowed	-0.22**	0.10	-0.18**	0.09	0.00	0.08	0.07	0.10
Immigrant	-0.28***	0.03	-0.26***	0.03	-0.34***	0.03	-0.34***	0.05
Partner 62+	0.09**	0.04	0.01	0.04	-0.04	0.02	0.05	0.03
F-test $\overline{z_i}$	358.02***				215.18***			
δ_{1t}	-1.40***	0.05			-1.09***	0.05		
δ_{2t}	-1.05***	0.05			-0.72***	0.05		
δ_{3t}	-0.20***	0.05			0.38***	0.05		
δ_{1t-m}	-1.32***	0.04			-0.09	0.07		
δ_{2t-m}	-0.97***	0.04			0.28***	0.07		
δ_{3t-m}	-0.15***	0.04			1.35***	0.07		
ρ_{tm}	0.95***	0.02			0.96***	0.03		
Obs.	21,547				20,515			
Chi^2	4,819				2,686			