Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Gradual retirement of men in the Netherlands

analysis of trends with an age-period-cohort model

J.L.M. Bonekamp

MSc Thesis 2015-035

Gradual retirement of men in the Netherlands:
analysis of trends with an age-period-cohort model

by

J.L.M. Bonekamp B.Sc.
(ANR: 195190)

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Econometrics and Mathematical Economics

Tilburg School of Economics and Management
Tilburg University

Supervised by:

Dr. J.A. Bolhaar
(CPB Netherlands Bureau for Economic Policy Analysis)

Prof. Dr. A.H.O. van Soest
(Tilburg University)

June 9, 2015

Abstract

This study investigates the (modest) increase in gradual retirement of men in the Netherlands. A multinomial logit age-period-cohort model is estimated for a sample of Dutch men born between 1921 and 1962, using data from the Dutch Labour Force Survey 1992-2012. The results indicate that the unobserved cohort effects are the most important factor in explaining the growth of gradual retirement for the period 1992-2012. There are, however, differences between the cohort effects of gradual retirement at an employer or gradual retirement through self-employment. Gradual retirement through self-employment did not become more prevalent over successive cohorts, as indicated by the unobserved cohort effects. Moreover, the unobserved age effects, without taking the option of gradual retirement through self-employment explicitly into account, suggest that, when possible, individuals prefer to reduce the number of hours worked towards the age when the individual is first eligible to an Old Age pension.

Acknowledgements

First of all, I would like to thank my supervisor at Tilburg University, Arthur van Soest, for all of the feedback (including correcting numerous incorrect uses of '-ly' and '-s'), discussions, and continuously pushing me to improve my thesis. Second, my supervisor at CPB Netherlands Bureau for Economic Policy Analysis (CPB), Jonneke Bolhaar, for making time to read through my thesis (and helping me to improve my style of writing), acting as a sounding board for most of my, sometimes completely 'out of the box', ideas, and for helping to make me feel welcome at CPB. Next, my other colleagues at CPB, and especially Daniel van Vuuren for some useful comments on an earlier draft of this thesis. Finally, I would like to thank my friends and family, for their insights and support.

Contents

Chapter 1

Introduction

The traditional retirement scenario, where an individual abruptly stops working at a given age, does no longer fit in a more recent view on the life-course trajectory (see, e.g., Brückner and Mayer (2005)). An alternative such as gradual retirement, where the individual gradually decreases the hours worked until he, or she, is fully retired, seems from the point of an individual intuitively more attractive (see, e.g., De Vaus, Wells, Kendig, and Quine (2007)). Gradual retirement can, in addition, serve as an instrument to facilitate individuals who do not yet want to fully retire to continue working part-time (either before or after the age when an individual is first eligible to an Old Age pension) thereby possibly allowing for an increase in the total labour supply of older individuals.

The main goal of this study is to investigate and disentangle the long term trends of gradual retirement of men in the Netherlands by an age-period-cohort analysis over the period 1992-2012. Moreover, do younger cohorts of men in the Netherlands have a larger probability to make use of gradual retirement programmes than older cohorts of men, holding age and period effects constant? How is the probability of a male individual choosing a gradual retirement program related to the age when an individual is first eligible to an Old Age pension in the Netherlands, holding period and cohort effects constant?

The current study attempts to address these questions through estimating a multinomial logit model for men aged 50-70 during the period 1992-2012 on the basis of a large, longitudinal survey: the Dutch Labour Force Survey. Two different identification strategies are conducted to disentangle gradual retirement into age, period, and cohort effects. The age effects includes life-cycle decisions such as timing of education, period effects include cyclical and instantaneously effects that affect everyone irrespective of the age, e.g. policy changes, and cohort effects[1] consists of the effects between

[1] Notice that there is some ambiguity whether some events, e.g. policy changes, should be classified as period or cohort effects.

different cohorts, e.g. the availability of certain retirement paths.

The outcomes of this study are of interest for several reasons. First, to get an idea of what to expect in the future concerning the long term trends of gradual retirement of men in the Netherlands. The modelling approach allows to identify factors that used to be important in the past, and whether these are still important nowadays. This approach also helps to identify factors that have become increasingly important over the years. Second, it could be of interest to the (Dutch) pension funds. Even if individuals prefer not to stop working abruptly, but prefer to gradually retire (see, e.g., Brückner and Mayer (2005)) it remains ambiguous whether an individual wants to retire before, or after, the age when the individual is first eligible to an Old Age pension (in the Netherlands[2]). Nevertheless, this is of interest for the pension funds in creating, for example, (personalized) pension products.

The results of the current study indicate that over successive cohorts gradual retirement of men in the Netherlands did become more prevalent. The unobserved cohort effects for individuals who gradually retire by working 1-34 hours increased approximately linear for individuals born between 1933-1950. For cohorts born in, or after, the 1950 the unobserved cohort effects still increased, but slower compared to cohorts born before the 1950s. The unobserved cohorts effects are similar for individuals who gradually retire by working 13-24 hours, or 25-34 hours. The unobserved cohort effects roughly linearly increase for individuals who gradually retire by working 1-12 hours for cohorts born between 1930-1960. The prevalence of gradual retirement through self-employment, however, hardly changed over successive cohorts born between 1930-1955. A decomposition of the growth in gradual retirement in the period 1992-2012 suggest that the growth can almost fully be explained by (unobserved) cohort effects. As for more recent cohorts these unobserved cohort effects hardly increases, it is unlikely that gradual retirement will become even more prevalent.

The estimated probability of gradual retirement by working 1-12 hours peaks at the age of 67, gradual retirement by working 13-24 hours peaks at the age of 62, and gradual retirement by working 25-34 hours peaks at the age of 58. This indicates that, when possible, individuals prefer to reduce the number of hours worked towards the age when they are first eligible to an Old Age pension.

The remainder of this study is organised as follows. Chapter 2 discusses (some of) the relevant literature on gradual retirement. Chapter 3 discusses the data en presents the empirical strategy. Chapter 4 discusses the results and provides some sensitivity analysis. Chapter 5 discusses the results, concludes, and provides some directions for future research.

[2]Until 2013 the age at which an individual was first eligible to an Old Age pension (in the Netherlands) used to be 65.

Chapter 2

Literature review

This section discusses some of the relevant literature on gradual retirement. First, different concepts of gradual retirement are explained. Next, the utilization of gradual retirement programs, and the effect of gradual retirement on total labour supply in the United States (from here on 'US'), Europe, and the Netherlands in particular, is examined. Finally, self-employment as a potential pathway of gradual retirement towards full retirement is discussed.

2.1 Gradual retirement

Gradual retirement, in contrast to traditional retirement, involves a time period during which gradual withdrawal from the labour force takes place until the individual is fully retired. To describe this transition, different authors use different concepts. This study uses mostly the same concepts as used by Scott (2004), Kantarci and van Soest (2008), Bloemen, Hochguertel, and Zweerink (2014) and others. Thus, gradual retirement is used to indicate a gradual withdrawal from the labour force. This could take place by, e.g., phased retirement (reducing the number of hours worked, but staying in the same job), or partial retirement (changing to a different, usually less demanding, job or becoming self-employed). Part-time retirement, where an individual earns part-time pension benefits and reduces the number of hours worked, is another concept of gradual retirement. Notice that part-time retirement does not necessarily involve a change of job, in contrast to partial retirement.

Other, in the literature commonly used, concepts of retirement transitions are 'bridge employment' and 're-entry'. Bridge employment is defined as the labour force status between career employment and full withdrawal of the labour force (Ruhm (1990); Zhan and Wang (2015)). The concept of re-entry, or "unretirement", is used for individuals that reverse their (early) retirement decision and re-enter the labour market (see, e.g., Maestas (2010)). This study will mostly focus on the concept of part-time retirement as a form of gradual retirement.

2.1.1 Utilization of gradual retirement in the US and Europe

In the US gradual retirement has become increasingly popular, as studies based on data from the Health and Retirement Study (HRS) conclude. Scott (2004) reports for data from the first fives waves (1992-2000) of the HRS that in 1998 about 18% was in either a phased or partial retirement program while in 1994 this was about 11%. A more recent study by Cahill, Giandrea, and Quinn (2015) focusses on the retirement patterns of three HRS cohorts: 'the Core' (respondents aged 51-61 in 1992), 'War Babies' (respondents born between 1942 and 1947), and 'Early Baby Boomers' (respondents born between 1948 and 1953). Their findings concerning phased or partial retirement of the first cohort correspond with the findings by Scott (2004). Furthermore, the authors conclude that the 'Early Baby Boomers' utilize the possibility of gradual retirement through bridge jobs (65% for men and 74% for women) even more than their predecessors did (HRS core: 57% and 54%)[1]. This increase is most likely induced by men and women who had to leave their career jobs involuntary.

In contrast to the increase in prevalence of gradual retirement in the US, Kantarci and van Soest (2008) conclude that their is no evidence that gradual retirement in Europe become more prevalent for the 1994-2000 period. Using data from the European Community Household Panel (ECHP) and the Panel Study of Income Dynamics (PSID) for workers aged 51-65 for the 1994-2000 period, they, in addition, concluded that the prevalence of gradual retirement varies substantially across different European countries. Their methodology is based on comparing part-time employment rates between the young and the old instead of using genuine data corresponding gradual retirement. Although it is likely that their findings hold for men, it is less likely that part-time employment is a good indicator for whether a woman is gradually retired (see, e.g., Been and Van Vliet (2014)).

In a more recent study Brunello and Langella (2013) corroborate that the prevalence of gradual retirement, at least for men, varies substantially across different European countries. Using data from the third wave (2008-2009) of the Survey on Health, Age and Retirement in Europe (SHARE), the authors find substantial differences between gradual retirement of men in the Central and Northern European countries[2] and the Mediterranean European countries[3]. The traditional pattern of retirement is more common in Mediterranean Europe than it is in Central and Northern Europe. In addi-

[1] The 'War Babies' utilized the possibility of gradual retirement through both bridge jobs (58% for men and 61% for women) and phased or partial retirement programs, more than their predecessors did (Giandrea, Cahill, & Quinn, 2009).

[2] Germany, Austria, Switzerland, Netherlands, Belgium, Denmark and Sweden.

[3] France, Italy and Spain.

tion, they conclude that gradual retirement is less common in Continental Europe than in the US.

2.1.2 Gradual retirement in the Netherlands

Bloemen et al. (2014) used Dutch administrative data on individuals aged 55-64 for the period 1999-2008 to investigate the current state of gradual retirement in the Netherlands for several industries. They report that the overall fraction of career job workers transitioning to partial retirement between ages 61 to 64 was 0.5% to 2% per year for men (0.5% to 1.5% per year for women), and the overall fraction of career job workers transitioning to phased retirement between ages 61 to 64 was 6% to 8% for men (9% to 12% for women). The remainder of their study tries to explain why this is less than the 34% that could potentially be achieved for the Netherlands according to Van Soest, Kapteyn, and Zissimopoulos (2007), and to provide some policy recommendations. Although the study by Bloemen et al. (2014) reports the 'current' state of gradual retirement in the Netherlands, they do not investigate, in contrast to this study, the long term trends of gradual retirement in the Netherlands. Moreover, they only report the current state of gradual retirement for individuals aged 55-64 for the period 1999-2008 whilst the group of gradual retired individuals aged 65-74 is of interest as well.

As touched upon by Kantarci and van Soest (2008), in order to investigate gradual retirement it would be possible to use the amount of part-time workers as a proxy variable. Most of the papers discussed up to this point only considered (part-time) workers up to the age of 65. One notable exception is a paper by the CPB Netherlands Bureau for Economic Policy Analysis (from here on CPB). They report that labour force participation (defined as (willing to) working for pay at least one hour a week) in the Netherlands of individuals aged 65-69 increased from 5.7% in 2001 to 12.7% in 2012 (CPB, 2014). It remains ambiguous whether part of this trend could be explained by a possible increase in gradual retirement.

2.1.3 Changes in fiscal facilitation of early full retirement

As of January 1, 2006, fiscal facilitation of early full retirement has changed for individuals born on or after January 1, 1950 due to changes in legislation. Before January 1, 2006, early retirement benefits were taxed while the early retirement premiums were exempted from taxation. Early retirement benefits are generally lower than the income of older working individuals, and, in addition, income of workers in the Netherlands is taxed progressively. Therefore, the marginal tax rate applying to early retirement contribution by workers is higher than the marginal tax rate of early retirement benefits. Thus, the fiscal facilitation before January 1, 2006, induced early full retirement, rather than gradual retirement for older individuals.

2.2 Gradual retirement and labour supply

Gradual retirement can affect the total labour supply in several ways. Without the possibility of gradual retirement, a worker could prefer working full-time over being fully retired. However, if the individual has access to a gradual retirement program, the individual could also prefer gradual retirement over working full-time thereby leading to a reduction in the labour supply of older individuals. On the other hand, if the individual has access to a gradual retirement program, the individual could also prefer gradual retirement over being fully retired thereby increasing the labour supply. Whether the positive or the negative labour supply effect dominates is of interest from a public policy point of view.

2.2.1 The US and Europe

Whether the positive labour supply effect dominates the negative labour supply effect remains rather ambiguous both internationally as in the Netherlands. Ghent, Allen, and Clark (2001) report that, using data from 15 campuses of the University of North Carolina, most of the faculty members would likely have continued working full-time if the phased retirement program was not introduced. The introduction of the phased retirement program led in this case, most likely, to a decrease of the labour supply. Wadensjö (2006), however, derives a different conclusion by performing a descriptive analysis of the abolishment of part-time pension schemes in Sweden. His research indicates that most of the workers that used to be eligible for a part-time pension after the abolishment rather fully retire, than continue working full time. Hence, in this case the part-time pension program most likely induced a positive labour supply effect.

The study by Ghent et al. (2001) and Wadensjö (2006) are micro-level studies that investigate the effect of gradual retirement on labour supply. Been and Van Vliet (2014), on the contrary, conduct a macro-level study using data from Eurostat for the period 1995-2008. For men, their findings suggest that part-time employment is a substitute for full early labour market withdrawal. Thus, for men, their results suggest that facilitating part-time work schemes may actually increase the labour supply at older ages.

Other studies focus on the working time constraints of older workers and the effect of these constraints on the labour supply of elderly individuals. Charles and DeCicca (2007) find for men, using multiple waves of the HRS, that these constraints encourage a transition from a career job towards full retirement rather than to gradual retire before becoming fully retired. Using data from the British Household Panel Survey for the period 1991-2004, Gielen (2009) reports that these time constraints mostly induce women to leave the labour market early. The results of both Charles and DeCicca (2007) and Gielen (2009) suggest that more flexibility by, e.g., gradual retirement could lead to an increase in labour supply of elderly individuals.

2.2.2 The Netherlands

In order to investigate whether gradual retirement plans could lead to an increase of total labour supply of elderly in the Netherlands, mostly stated-preferences studies are conducted (Bruinshoofd and Grob (2006); Van Soest et al. (2007); Kerkhofs, Fouarge, and Ester (2009); Van Soest and Vonkova (2014)). A stated-preference approach is generally used to gather knowledge, by describing hypothetical situations in a survey, about the preferences of individuals concerning products that are not (yet) widely available. Hence, to investigate the preferences of different (gradual) retirement paths a stated-preference approach could be convenient. Van Soest and Vonkova (2014) used stated-preference data from a questionnaire fielded in 2006, 2007 and 2008 to participants of the Dutch CentERpanel[4]. The authors conclude that the estimated retirement age of an individual of the investigated (gradual) retirement paths are sensitive to financial incentives such as the replacement rate. The financial incentives determine whether facilitating gradual retirement would possibly result in an increase in total labour supply of older individuals.

Fouarge, De Grip, and Montizaan (2012) report that facilitating a gradual retirement program would lead in the Netherlands to older workers postponing their retirement. However, prior to their full retirement the older individuals would reduce the number of hours worked. As the positive effect on the labour supply and the negative effect cancel out, facilitating a gradual retirement program would not affect the total labour supply of older workers in the Netherlands. Using responses of a survey of participants in the CentERpanel in 2005, a stated-preference experiment was conducted to obtain their results.

A study that reports a negative total labour supply effect in the Netherlands is conducted by Vroom, Van Sonsbeek, and Albas (2012). The authors estimate a transitional model using data of Statistics Netherlands (CBS) for the period 2005-2008, and use microsimulations to conclude that gradual retirement leads to a decrease of the total labour supply of older people in the Netherlands. However, due to lack of data, the possibility of individuals working after the age of 65 is not considered. Moreover, in their analysis they do not include individuals that are self-employed. Nevertheless, self-employment could be an attractive alternative if phased or partial retirement is not possible, or promoted, by the individuals employer (Bloemen et al.

[4]CentERpanel is a representative panel of the Dutch adult population.

(2014)).

To summarize, several studies provide conflicting results of the hypothesis that gradual retirement leads to a positive total labour supply effect for older individuals in the Netherlands. I will conduct an age-period-cohort analysis, rather than a stated-preference approach or a transitional model, using data from the Dutch Labour Force Survey for the period 1992-2012 to disentangle the long term trends of gradual retirement in the Netherlands. This may help to determine whether gradual retirement in the Netherlands up to this point lead to an increase, or decrease, in total labour supply of older individuals.

2.3 Self-employment as an alternative gradual retirement path

The amount of research across European countries on labour participation of older individuals that allows for the possibility of part-time and self-employment is rather limited. An exception to this, is a study by Been and Van Vliet (2014) where the variation in withdrawal of older worker across European countries over the period 1995-2009 is analysed. They conclude that for workers aged 55 to 64 self-employment does not act as a substitute for gradual retirement, but rather to end unemployment. Studies that use the HRS, however, find conflicting results. Zissimopoulos and Karoly (2009) argue that many workers during the 1992 to 2002 interval, transitioned to self-employment after the age of 50 as part of transition to full retirement. Moreover, Giandrea, Cahill, and Quinn (2013) find that a substantial part (21% for men and 10% for women) of workers aged 51 to 61 already in 1992 was self-employed. In 2010 this rate nearly doubled (41% for men and 22% for women) for the same group of individuals. The authors argue that this increase is mainly the result of self-employed individuals who retire later compared to wage-and-salary workers, and more wage-and-salary workers who switch into self-employment later in life than vice versa.

Also in Europe, and thus in the Netherlands as well, self-employed individuals retire later than non self-employed individuals (Hochguertel, 2010). Bloemen et al. (2014) argue that a possible explanation for the higher effective retirement age may be the result of better opportunities of gradual retirement when self-employed, as in the Netherlands, due to labour agreements, it is difficult or even impossible to continue to work in a paid job beyond the age of 65. Nevertheless, Been and Knoef (2015) indicate that towards the end of a career self-employment is used to end or avoid a period of unemployment, instead of an alternative gradual retirement route. That is, for individuals between the ages of 50 and 63 thereby not considering gradual retirement after the age of when an individual is first eligible to an Old Age pension in the Netherlands. Their results would probably have been different if they would have been able to include individuals aged 64 and older. As the methodology and the data in the current study allows to distinguish the effects for individuals that are self-employed and gradually retired even after the age of 65, this is included in the analysis.

Chapter 3

Methodology

This chapter presents a description of the data and the definition in which cases this studies identifies an individual as gradually retired. In addition, the empirical strategy of this study is discussed.

3.1 Data description

This study uses data from the Dutch Labour Force Survey (DLFS) 1992-2012. The DLFS is a survey conducted among residents of the Netherlands, with the exception of individuals living in institutions like nursing homes or prisons. Every year, a new random sample of 1% of the Dutch population aged 15 year or older is interviewed. Thus, multiple observations for a given individual are not observed. The DLFS collects information on household's statistics and the individual labour market situation.

Although the data of the DLFS contains a lot of information, there is no explicit information available on the (gradual) retirement status of an individual. Thus, in order to disentangle the long term trends of gradual retirement, a proxy variable is necessary. Kantarci and van Soest (2008) mentioned that part-time employment indicates how many older workers withdraw from the labour market. In addition, Been and Van Vliet (2014) conclude that "men use part-time employment as a step in gradual transition from full-time employment to retirement." For women, however, part-time employment is less likely to capture the movements of gradual retirement. In contrast to men, women, especially in the Netherlands, tend to work part-time a substantial part of their career and a change in number of hours worked towards retirement is less imminent (Bosch, Deelen, and Euwals (2010); Euwals, Knoef, and Van Vuuren (2011)). Therefore, the focus for the remainder of this study will be on men.

Table 3.1: Summary statistics, men aged 50-70, in percentages or years.[a]

	Fully working	Partly retired	Fully retired
Self-employed	19.41	24.75	0.00
Position in household			
Cohabiting	89.32	85.37	84.10
Single	9.73	13.57	14.66
Other	0.95	1.06	1.23
Children[b]			
No children	51.19	69.51	85.31
Only children < 18	12.45	9.30	2.70
Both children < 18 and >=18	9.62	4.61	1.34
Only children >= 18	26.74	16.28	10.65
Education[c]			
Primary	7.60	7.54	16.40
Lower secondary	19.48	16.48	23.18
Higher secondary	40.05	35.92	38.44
Tertiary	32.87	40.06	22.98
Age	54.98	58.42	62.80
Cohort (year of birth)	1951	1948	1941
VPL	63.13	41.91	10.82
Number of observations	350,187	93,304	259,388

Source: DLFS, 1992-2012

[a] Weighted summary statistics.

[b] 'Children' refers to children in the household.

[c] The education levels are defined as follows: 'Primary' indicates that no secondary education is completed (just primary school); 'Lower secondary' indicates that lower vocational or general school is completed (in Dutch: VMBO); 'Upper secondary' indicates that advanced vocational or general school is completed (in Dutch: MBO, HAVO, VWO); 'Tertiary' indicates that academic or vocational colleges is completed (in Dutch: HBO, WO).

Figure 3.1: Rate of men working part-time by age (in %). *Note:* Cohorts in 3-year groups, from cohort born in 1930-1932 to cohort born in 1948-1950. *Source:* Own calculations based on data from the DLFS, 1992-2012.

Gradual retirement: definition in this study

An individual is defined as working part-time, and thus gradually retired, if he works strictly less than 35 hours per week. This definition of part-time employment is in correspondence with the definition as used by CBS and the US Bureau of Labor Statistics. In addition, I assume that an individual can either be fully retired, gradually retired, or working full-time. Individuals who are unemployed, but willing to work at least one hour per week, are, as a result of this definition, not included in the analysis of this study. As part of the sensitivity analysis other definitions of gradual retirement will also be investigated.

The observed rates of men working part-time by cohort and age are shown in Figure 3.1. Especially for the more recent cohorts the rate of men working part-time starts to increase from the age of 50 onwards until the age of 63 and tends to decrease afterwards. At a given age it can be observed from the overlapping segments that jumps of a few percent are not uncommon. These ‘jumps’ are combined period and cohort effects. Since the increase of men working part-time starts before the age of 55, this study will, in contrast to, e.g. Bloemen et al. (2014), Vroom et al. (2012), and Van Soest and Vonkova (2014), also include men aged 50-55 for this analysis. A closer inspection of the data indicates that for most of the periods after the age of 70 over 90% of the men is fully retired. In addition, the number of individuals that is not fully retired is insufficient to obtain significantly relevant estimates, therefore the sample is restricted to male individuals up to and including age 70.

Individuals are grouped together in a cohort by birth year. To ensure that each cohort consists of at least 5,000 observations, individuals born in the 1921-1926 period form a single cohort, individuals born in the 1927-1929 period form one cohort, individuals born in the 1930-1932 period are given the same cohort indicator, and also individuals born in the 1961-1962 period form a single cohort. In addition, to capture the effects of the changes in legislation as discussed in Section 2.1.3 a dummy variable, named VPL, is constructed which equals 1 if the individual is born in, or after, 1950 and zero otherwise.

The resulting dataset contains 702,879 observations for men. The oldest cohort was born in 1921 and the youngest cohort in 1961. Furthermore, the smallest category ('Other') still counts in total for more than 7,500 observations (Table 3.1). On average, almost 66% have no children living at home.

3.2 Empirical Strategy

This section presents the empirical strategy of this study. First the methodology of a multinomial logit model is presented. Next, the methodology of an ordered logit model is discussed. Furthermore, this section will discuss how this study tackles the identification problem of an age-period-cohort analysis, the strategy to relax the independence of irrelevant alternatives assumption in this study, and finally some methods to relax the distributional assumptions are presented.

3.2.1 Multinomial logit with J alternatives

This study initially assumes that an individual faces three possible choices: he either works full-time, retires gradually, or is fully retired[1]. While one individual could prefer working full-time over gradually retiring, another individual could prefer gradual retirement over working full-time. As there is not an obvious ranking of each of the alternatives, a multinomial logit (MNL) model is most appropriate. A MNL model is in correspondence with most of the studies that model gradual retirement decisions (see, e.g., Thomson (2007); Cahill et al. (2015); Bloemen et al. (2014)).

To formalize, at a given time there are J alternatives, indexed $j = 1, \ldots, J$. The utility that individual i attaches to each of these alternatives is given by u_{ij}. Individual i will choose the alternative which provides the most utility, that is, if $u_{ij} = \max(u_{i1}, \ldots, u_{iJ})$. Since these utilities are not observed additional assumptions are necessary. Therefore, assume that

[1]The DLFS does not contain information of the individual in the period previously to the survey. Therefore, the option of "unretirement" is not taken into account.

$u_{ij} = \mu_{ij} + \epsilon_{ij}$ where μ_{ij} is a non stochastic function of observables and ϵ_{ij} are assumed to be mutually independent Type I extreme value distributed[2]. Then it follows, see McFadden (1973), that

$$\begin{aligned} P(y_i = j) &= P(u_{ij} = \max(u_{i1}, \ldots, u_{iJ})) \\ &= P(\mu_{ij} + \epsilon_{ij} > \max_{k=1,\ldots,J,k\neq j}(\mu_{ij} + \epsilon_{ij})) \\ &= \frac{\exp(\mu_{ij})}{\sum_{k=1}^{J} \exp(\mu_{ik})} \end{aligned}$$

where $y_i = 1, \ldots, J$. By specifying the distribution of ϵ_{ij} the scaling of the utility is set. In order to set the location of the utility one usually normalizes one of the deterministic utility levels to zero, so $\mu_{i1} = 0$ (see, e.g., Verbeek (2012)). Appendix B tests the Independence of Irrelevant Alternatives (IIA) assumption of this model for $J = 3$ and $J = 5$ fitted with the data of the DLFS.

To incorporate both observed individual characteristics and age, period and cohort effects, μ_{ij} is defined as:

$$\mu_{ij} = \beta_{0,j} + \beta_{1,j} x_{it} + g_a(a_{it}|\theta_{a,j}) + g_t(t|\theta_{t,j}) + g_c(c_i|\theta_{c,j}) \tag{3.1}$$

where x_{it} is a vector of variables including individual and family characteristics, a_{it}, t, and c_i respectively denote age, period, and cohort effects, and corresponding transformation functions are denoted by g. $\theta_j = (\beta_{0,j}, \beta_{1,j}, \beta_{a,j}, \beta_{t,j}, \beta_{c,j})$ is a vector that contains the parameters. Thus, the probability of an individual i choosing alternative j at a given period becomes:

$$\begin{aligned} &P(y_i = j|x_{it}, a_{it}, t, c_i) = \\ &\frac{\exp(\beta_{0,j} + \beta_{1,j} x_{it} + g_a(a_{it}|\theta_{a,j}) + g_t(t|\theta_{t,j}) + g_c(c_i|\theta_{c,j}))}{1 + \sum_{k=2}^{J} \exp(\beta_{0,k} + \beta_{1,k} x_{it} + g_a(a_{it}|\theta_{a,k}) + g_t(t|\theta_{t,k}) + g_c(c_i|\theta_{c,k})} \end{aligned} \tag{3.2}$$

The marginal effect of variable x_t is computed (see, e.g., Cameron and Trivedi (2005) p. 502), denoting the probability in (3.2) by p_{jt} (thus omitting the individual subscript), by

$$\frac{\partial p_{jt}}{\partial x_t} = p_{jt}(\beta_j - \bar{\beta}),$$

where $\bar{\beta} = \sum_l p_{lt}\beta_l$ is a probability weighted average of β_l. Notice that the sign of the marginal effect is not of the same sign of β_j, unless $\beta_j > \beta_k$ for all $k \neq j$.

[2] The cumulative distribution function of each ϵ_{ij} is given by $F(t) = \exp(-\exp(-t))$

3.2.2 Ordered logit with J alternatives

In contrast to the assumption of the previous section, there may exists a natural ordering for an individual to be fully working, gradually retired or fully retired. An ordered logit/probit, model would then be more appropriate for the analysis.

To formalize, at a given time there are J alternatives, indexed $j = 1, ..., J$, and assume that the underlying process is characterized by

$$y_i^* = x_i'\beta + \epsilon_i$$

where y_i^* is a latent variable. By specifying the distribution of ϵ_i the scaling normalization is set. For the ordered logit model ϵ is logistic distributed, and for the ordered probit model ϵ is standard normal distributed. To set the location normalization, assume that x does not include an intercept (see, e.g., Cameron and Trivedi (2005) p. 519).

To incorporate both observed individual characteristics and age, period and cohort effects, y_i^* is defined as:

$$y_i^* = \beta_1 x_{it} + g_a(a_{it}|\theta_a) + g_t(t|\theta_t) + g_c(c_i|\theta_c) + \epsilon_i$$

where the the definition of the variables are in correspondence with the previous section. $\theta = (\beta_1, \beta_a, \beta_t, \beta_c)$ is a vector that contains the parameters. An individual chooses alternative j if y_i^* is in between two (unknown) thresholds. Thus,

$$y_i = j \qquad \text{if} \quad \alpha_{j-1} < y_i^* \leq \alpha_j \tag{3.3}$$

where $\alpha_0 = -\infty$ and $\alpha_J = \infty$. The probability to choose alternative j is defined (after some rewriting) as

$$\begin{aligned} P(y_i = j|x_{it}, a_{it}, t, c_i) = \\ &F(\alpha_j - (\beta_1 x_{it} + g_a(a_{it}|\theta_a) + g_t(t|\theta_t) + g_c(c_i|\theta_c))) \\ &\quad - F(\alpha_{j-1} - (\beta_1 x_{it} + g_a(a_{it}|\theta_a) + g_t(t|\theta_t) + g_c(c_i|\theta_c))), \end{aligned}$$

where F is the cumulative distribution function of ϵ.

The Proportionality assumption

The Proportionality assumption (for the ordered logit model sometimes referred to as the 'proportional odds assumption') is essential to a valid application of the ordered logit/probit model. The proportional odds assumption indicates that the logarithm of the odds forms an arithmetic sequence. Under the assumption that ϵ is standard logistically distributed, the cumulative distribution probabilities $\gamma_{i,j} = P(y_i \leq j|\mu_i(x_{it}, a_{it}, t, c_i))$ can be written (see, e.g., Brant (1990)) as

$$\log(\gamma_{i,j}/(1 - \gamma_{i,j})) = \alpha_j - \theta'\mu_i. \tag{3.4}$$

Thus, the parameter estimates, θ, are identical for each category.

In Appendix C the Proportionality assumption of this model for both $J = 3$ and $J = 5$, fitted with the data of the DLFS, is tested. As the test(s) reject the null hypothesis of the Proportionality assumption to hold in favour of the alternative hypothesis, the remainder of this study focusses on the results of the MNL estimates.

3.2.3 Tackling the identification problem in APC analysis

The key problem in age-period-cohort (APC) analysis using Model (3.1) and Model (3.3) is that not all parameters can be identified when all of the transformation functions g_a, g_t, and g_t contain a linear term. If the period and age of an individual is known, the cohort of this individual is known by the following equation:

$$t - a_{it} = c_i,$$

thus a perfect linear relationship. An extensive literature going back to the 1970s examined the identification problem in the APC analysis and has identified three conventional strategies for identification: assuming that the two most recent cohorts were identical (Mason, Mason, Winsborough, and Poole (1973); Fienberg and Mason (1978)), using a "proxy" variable approach that assumes that the cohort or period effects are proportional to certain measured variables (Heckman and Robb (1985); O Brien (2000)), and defining a functional form for at least one of the age, period, or cohort variables such that the relationship becomes non-linear (Mason et al. (1973)).

The difference in peaks at a given age of Figure 3.1 indicates that the choice of the approach to tackle the identification problem of the APC analysis could largely determine the estimates of the age, period and cohort effects. Implementing both the 'functional form approach', and the 'proxy variable approach' provides an opportunity to test this conjecture. In order to avoid arbitrary results it is necessary to support the assumptions necessary for identification by, e.g., economic theory.

Functional form approach

In the Netherlands, older cohorts used to have access to both official and unofficial, beneficial paths towards (gradual) retirement (Kerkhofs, Lindeboom, and Theeuwes (1999)). Through reforms by the Dutch government, these unofficial retirement paths by disability insurance or unemployment became less beneficial or infeasible for more recent cohorts (Euwals, Van Vuren, and Van Vuuren (2011)). Moreover, by reforming the actuarially unfair early retirement schemes in the Netherlands the decision of the more recent cohorts to (gradually) retire is delayed compared to the older cohorts (Euwals, van Vuuren, and Wolthoff (2010)). Thus, these reforms have quite an impact on the cohort effects to (gradually) retire. However, as suggested by the literature, these cohort effects diminish over time. Therefore, this study assumes as a 'functional form' that these cohort effects follow a reversed "S" shape over time. Thus, in practice this means that $g_c(c_i|\theta_{c,j})) = -\theta_{c,j}/(1 + \exp(-c_i))$.

Proxy variable approach

A macroeconomic factor is mostly used as a proxy variable that is assumed to be proportionally to the period effects (see, e.g., Kapteyn, Alessie, and Lusardi (2005); Euwals, Knoef, and Van Vuuren (2011)) in economic APC analysis. A macroeconomic factor that influences the timing of (gradual) retirement is, e.g., the unemployment rate. A weak labour market affects the workers perception of their future job prospects and could thus induce, as a result of the definition of gradual retirement in this study, both (gradual) retirement and involuntary gradual retirement (Fischer and Sousa-Poza (2006); Dorn and Sousa-Poza (2010); Bosworth and Burtless (2010); Gorodnichenko, Song, and Dmitriy (2013)).

The APC analysis, using the proxy variable approach to overcome the identification problem, is conducted with the unemployment rate as proxy variable. It is assumed that the period effects are only determined by this macroeconomic indicator. However, notice that not all relevant period effects may be captured by the proxy variable approach and thus care is necessary when interpreting the estimation results. The annual aggregated unemployment level by education group (U_t) in the Netherlands for men aged 15-64, as published by CBS, at time t is used (thus, $g_t(t|\theta_{t,j}) = \theta_{t,j}U_t$).

3.2.4 Relaxing the IIA assumption

The MNL model is characterized by the IIA assumption which sometimes is too restrictive. To generalize the the MNL model, thereby relaxing the IIA assumption, a nested logit (NL), a random parameter logit (RPL) model, or a multinomial probit (MNP) are often applied to the data.

A general discussion of the NL model can be found in, e.g., McFadden (1978). The idea of the NL model is to break the decision making process into groups. At the top level there are some limbs to choose from, and each limb is assumed to have some branches. The errors are allowed to be correlated in each limb but not between limbs. To estimate a NL model (see, e.g., Cameron and Trivedi (2005) p. 510) a set of regressors that pertain only to the limb and a set of regressors that pertain to both limb and branch is necessary. As the DLFS lacks variables that distinguish between alternatives (limbs), estimation of the NL model is not feasible in this study.

Alternatively, a RPL model (see, e.g, Revelt and Train (1998)) could be estimated. However, the same individual should be observed in the data multiple times to consistently account for the taste heterogeneity patterns (see, e.g., Rose, Hess, Bliemer, and Daly (2011)). Another way to induce correlation in the unobserved component between each alternative is to work with normally distributed error terms. The resulting model is known as the Multinomial probit (MNP) (see, e.g., Daganzo (1979)).

Multinomial probit with three alternatives

This section assumes that an individual faces three possible choices: to work full-time, to gradually retire, or to fully retire. Thus, at a given time there are three alternatives, indexed $j = 1, 2, 3$. The utility that individual i attaches to each of these alternatives is given by

$$\begin{aligned} u_{i1} &= \mu_i' \theta_1 + \epsilon_{i1} \\ u_{i2} &= \mu_i' \theta_2 + \epsilon_{i2} \\ u_{i3} &= \mu_i' \theta_3 + \epsilon_{i3} \end{aligned}$$

where $\mu_i(x_{it}, a_{it}, t, c_i)$ is a matrix of observables for individual i (thus, non stochastic), and θ_j is a vector that contains the parameters. In contrast to Section 3.2.1, where the stochastic part is assumed to be mutually independent with a Type I extreme value distribution, the stochastic part is here assumed to be multivariate normally distributed with covariance matrix

$$\text{cov}\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}.$$

Following Dansie (1985), Keane (1992), and Train (2009) identification is achieved by normalizing the utility of one of the alternatives to zero and restricting one covariance matrix element. Setting $u_{i1} = 0$ and $\sigma_{22} = 1$, the model

$$\begin{aligned} u_{i1} &= 0 \\ u_{i2} &= \mu_i' \theta_2 + \epsilon_{i2} \\ u_{i3} &= \mu_i' \theta_3 + \epsilon_{i3} \end{aligned}$$

with error covariance matrix

$$\text{cov}\begin{bmatrix} \epsilon_2 \\ \epsilon_3 \end{bmatrix} = \begin{bmatrix} 1 & \sigma_{23} \\ \sigma_{23} & \sigma_{33} \end{bmatrix}.$$

is obtained.

The individual chooses the alternative with the highest utility. For notational purposes define $V_{ij} = \mu_i'\theta_j$, then the probability to choose the second alternative becomes

$$\begin{aligned} P(y_i = 2) &= P(u_{i2} > u_{i1}, u_{i2} > u_{i3}) \\ &= P(u_{i2} > 0, u_{i2} - u_{i3} > 0) \\ &= P(V_{i2} + \epsilon_{i2}, V_{i2} - V_{i3} + \epsilon_{i2} - \epsilon_{i3} > 0) \\ &= P(\epsilon_{i2} > -V_{i2}, \epsilon_{i2} - \epsilon_{i3} > -(V_{i2} - V_{i3})) \\ &= \int_{-(V_{i2}-V_{i3})}^{\infty} \int_{-V_{i2}}^{\infty} f(\epsilon_{i2}, \epsilon_{i2} - \epsilon_{i3}) d(\epsilon_{i2}) d(\epsilon_{i2} - \epsilon_{i3}), \end{aligned}$$

which is a bivariate integral that does not have a closed-form solution. This, and analogous expressions for the other observed probabilities, forms the basis of the log-likelihood function. Let

$$y_{ij} = 1(y_i = j),$$

where $1(I)$ is the indicator function which equals one if I is true and zero otherwise. The log-likelihood function is defined by:

$$\mathscr{L}(\sigma_{23}, \sigma_{33}, \theta_2, \theta_3) = \sum_{i=1}^{n} \sum_{j=1}^{3} y_{ij} \log(P(y_i = j)) \tag{3.5}$$

where the total number of observations in the sample is n.

3.2.5 Relaxing the distributional assumption

The previous section discussed several models that did not have the IIA assumption. Most of these models, however, did (sometimes implicitly) assume some sort of distributional form of the error term. This section presents some strategies to relax the distributional assumption on the stochastic component of a multinomial 'logit' model[3].

As non-parametric methods suffer from the so called "curse of dimensionality", only semi-parametric single index methods where there is only one non-parametric dimension are investigated. These semi-parametric multinomial logit models are often estimated by a penalized likelihood approach (see, e.g., Kneib, Baumgartner, and Steiner (2007)), or, more recently, using Bayesian econometrics (see, e.g., Li (2011)). These models require alternative specific variables which the DLFS is without. In addition, these models fail to relax the IIA assumption (although the authors sometimes present a mixed model based extension to relax the IIA assumption). Next to requirement of alternative specific variables, at least one variable is required

[3]Notice that under the assumption of independent and identically distributed error terms the estimated coefficient by the multinomial logit model and the multinomial probit are comparable (see Stern (1989)).

to be continuous. Although one could argue that the variable age is approximately continuous, the APC methodology of this study requires that age is implemented in the model as a set of binary variables.

The semi-parametric binary response model by Klein and Spady (1993), based on kernel density estimation, does not require a continuous variable or alternative specific variables. Generalizing their methodology may thus lead to a, for this study, feasible methodology to relax the distribution assumption on the error term. Lee (1995) generalized the methodology of Klein and Spady (1993), but his methodology requires at least one continuous variable. Thus, as these models are not suited for this study to relax the distributional assumption on the error term, this study will only provide results under some sort of distributional assumption on the error term.

Chapter 4

Results

This chapter presents and discusses the estimation results obtained by the models of Section 3.2. First, the estimation results of the APC analysis by a MNL model where an individual can either be fully working, gradually retired, or fully retired are reported. Furthermore, the estimation results where an individual could be gradually retired through self-employment, or as an employee, are presented. Next, the estimation results are shown when gradual retirement is divided into three categories: working 1-12 hours, working 13-24 hours, and working 25-34 hours. These estimation results are derived via a MNL model. Also the estimation results where an individual can be fully working, gradually retired, or fully retired derived under a MNP model are discussed. Finally, the effect of (possible) misclassification of gradual retirement is presented.

4.1 MNL: fully working, gradually retired, or fully retired

This section presents the estimation results of the APC analysis by a MNL model where an individual faces three possible retirement options: fully working, gradually retired, or fully retired. The marginal effects, the change in probability of gradual retirement (or the change in probability to fully retire) due to a change in a certain exogenous variable under the proxy variable approach are reported in Table 4.1. Also Table 4.2 presents the marginal effects, but under the functional form approach. The effects of partly retired and fully retired add up to the marginal effects on the probability of fully working (with a plus or minus sign).

4.1.1 Under the proxy variable approach

Being single increases both the probability of gradual retirement, and the probability of full retirement. Having children living at home decreases the

Table 4.1: Marginal effects for the probability of partly retired or fully retired (in %-points), men aged 50-70 under the proxy variable approach.[a,b,c,d]

	Partly retired		Fully retired	
	Marginal effect	t-value	Marginal effect	t-value
Position in household				
Cohabiting				
Single	1.86	15.48	7.70	60.75
Other	2.56	6.54	6.09	15.23
Children				
No children				
Only children < 18	2.89	18.42	−4.28	−20.79
Both children < 18 and >=18	−0.16	−0.78	−7.04	−25.90
Only children >= 18	−1.29	−11.33	−5.20	−42.03
Education				
Primary				
Lower secondary	0.15	0.70	−7.52	−36.22
Higher secondary	1.39	5.47	−9.98	−41.05
Tertiary	5.82	20.98	−15.96	−59.72
Age				
Dummy variables	Yes	***	Yes	***
Period (year)				
U_t	−0.17	−4.56	−0.02	−0.56
Cohort (year of birth)				
Dummy variables	Yes	***	Yes	***
VPL	13.11	14.13	−20.70	−21.49

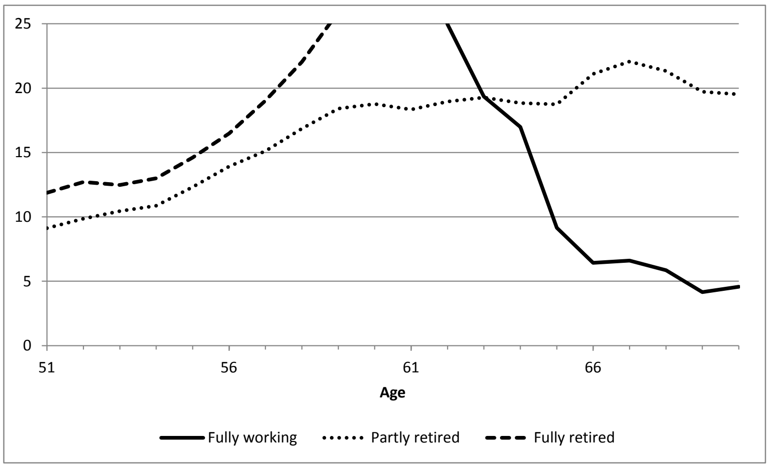
[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

[b] The model for the probability of fully working or partly retired is estimated by a multinomial logit model with fully working as the reference category (see Table A.1 for the estimates of this model).

[c] The age effects are presented in Figure 4.1, the cohort effects in Figure 4.2 and the corresponding marginal effects can be found in Table A.3.

[d] The reported point estimates are average marginal effects. That is, for each individual the marginal effect is calculated and averaged afterwards.

(a)



(b)

Figure 4.1: (a) Probability of working full-time, partly retired or fully retired by age (in %) under the proxy variable approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the age dummies. (b) As for (a) but for a different scaling of the y-axis.

probability of full retirement, while the effect on the probability of gradual retirement depends on the age of children living at home. The marginal effect is negative when children of at least 18 years old are living at home, while the marginal effect is positive when there are children younger than 18 years involved. Furthermore, educational attainment increases the probability of gradual retirement while it decreases the probability of full retirement.

One of the interests of this study are the age effects on the choice of gradual retirement (for the estimated profile of the age effects, see Figure 4.1). The age effects should be interpreted as unobserved age effects, as

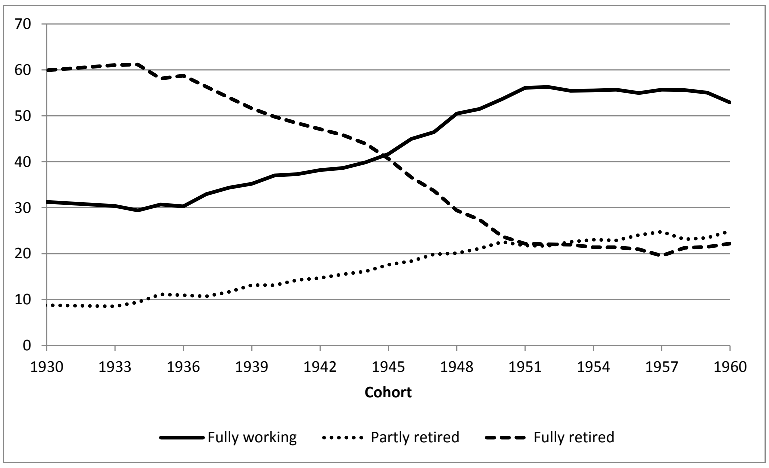


Figure 4.2: Probability of working full-time, partly retired or fully retired by cohort (in %) under the proxy variable approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the cohort dummies (which include the VPL dummy).

observed variables such as educational attainment will contain some age effects as well. The results show a gradual increase in the probability of becoming gradually retired by working 1-34 hours until the age of 60 after which it remains constant until the age of 65. The preferences towards gradual retirement by working a certain number of hours could change when an individual becomes older, this hypothesis will be addressed in Section 4.3. After the age of 65 the probability of gradual retirement increases again for a few years. The second increase may be (partly) the result of individuals who gradually retire through self-employment. This conjecture will be investigated in Section 4.2.

The cohort effects[1] on the choice of gradual retirement of Figure 4.2 show roughly a linearly increase starting from individuals born in 1933 until the cohort of 1951. The probability of the choice of fully working and fully retired are roughly constant for the oldest cohort after which the probability of the choice of full retirement decreases linear until the cohort of 1951, while the probability of the choice of fully working increases linearly in the same period. After the cohort of 1951 both the probability of choosing full retirement and the probability of fully working remain roughly constant, whereas the probability of gradual retirement tends to increase.

Part of the (minor) increase in the probability of gradual retirement for individuals born after the 1950s could possibly be explained by an increase in the availability or awareness of gradual retirement plans. Alternatively, it could be the result of changes in social norms towards gradual retirement.

[1] Similar to the age effects, the cohort effects should be interpreted as unobserved cohort effects.

The substantial decrease in the probability of full retirement starts for individuals born in 1933. These individuals were first affected by changes in fiscal facilitation of early retirement (see Kerkhofs et al. (2009)). As of January 1, 2006, also these fiscal facilitation of early retirement changed, affecting individuals born on or after January 1, 1950 (see Section 2.1.3 of this study). Thus, part of the changes in the (unobserved) cohort effects could be initiated by changes in legislation to which the individuals respond.

4.1.2 Under the functional form approach

The estimated marginal effects of Table 4.2 are mostly in correspondence with the marginal effects of Table 4.1, except for the estimated marginal effect of educational attainment on the probability to gradually retire. The marginal effects of Table 4.2 on educational attainment are on average 0.79 percentage point higher. In addition, the age effect on the choice of gradual retirement under the functional form (see Figure 4.3) behaves similar to the age effects on the choice of gradual retirement under the proxy variable, except for the ages between 59-65. Instead of remaining constant, Figure 4.3 shows that the probability to gradually retire decreases.

The different estimation results under the proxy variable approach and under the function form approach are most likely induced by an, for this study, undesirable effect of the functional form approach. The functional form approach implicitly assumes, in the setting of this study, that the cohort effects are identical for all individuals (irrespective of, e.g., educational attainment). However, this is not necessarily true as, among others, the labour force participation of older men in the Netherlands differs between lower and higher educated individuals (Van Vuuren and Deelen (2009); Van Vuuren (2014)). The proxy variable approach provides a more flexible and elegant approach to tackle the identification problem inherent to an APC analysis. Therefore, the remainder of study solely focusses on the APC analysis under the proxy variable approach.

4.1.3 Sensitivity analysis and underlying assumptions

If the age effects and cohort effects are separable from the other explanatory variables, the age effects and cohort effects can be interpreted as the trend applicable to all men irrespective of their individual characteristics. To test this conjecture some sensitivity analysis is conducted. In addition, this section investigates the effect of downwards adjusting the cut-off point in the definition of gradual retirement. The results are presented under the proxy variable approach.

If the lines of Figure 4.4 would run (roughly) parallel to each other it would support the assumption that the age and cohort effects are separable

Table 4.2: Marginal effects for the probability of fully working or partly retired (in %-points), men aged 50-70 under the functional form approach.[a,b,c,d]

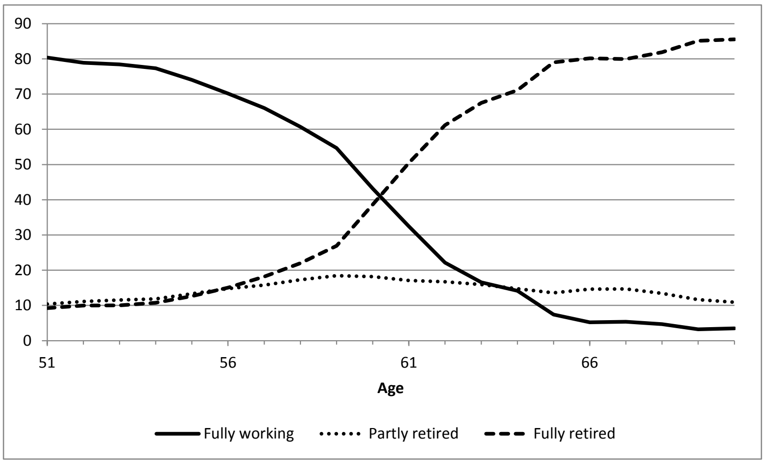
	Parlty retired		Fully retired	
	Marginal effect	t-value	Marginal effect	t-value
Position in household				
Cohabiting				
Single	1.90	15.79	7.66	60.63
Other	2.60	6.64	6.05	15.15
Children				
No children				
Only children < 18	2.91	18.59	−4.26	−20.73
Both children < 18 and >=18	−0.13	−0.64	−7.00	−25.79
Only children >= 18	−1.20	−10.53	−5.17	−41.79
Education				
Primary				
Lower secondary	0.71	4.12	−7.44	−47.95
Higher secondary	2.22	14.12	−9.87	−69.83
Tertiary	6.80	43.52	−15.85	−107.47
Age				
Dummy variables	Yes	***	Yes	***
Period (year)				
Dummy variables	Yes	***	Yes	***
Cohort (year of birth)				
1/(1+exp(-(cohort-1945)))	1.34	6.11	−8.30	−37.72

[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

[b] The model for the probability of fully working or partly retired is estimated by a multinomial logit model with fully working as the reference category (see Table A.4 for the estimates of this model).

[c] The results corresponding age effects are presented in Figure 4.3, and the corresponding marginal effects can be found in Table A.6.

[d] The reported point estimates are average marginal effects. That is, for each individual the marginal effect is calculated and averaged afterwards.



(a)

(b)

Figure 4.3: (a) Probability of working full-time, partly retired or fully retired by period (in %) under the functional form approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the age dummies. (b) As for (a) but for a different scaling of the y-axis.

from the other explanatory variables. Only the impact of the position in the household and the cohort effects seems to be separable (Figure 4.4, upper right). Consequently, when interpreting the age and cohort effects (some of) the individual characteristics should be taken into account.

The DLFS lacks a variable that indicates whether an individual is gradually retired. To overcome this problem, a proxy variable is introduced in Section 3.1 to indicate whether an individual is gradually retired. Consequently, the age and cohort effects of gradual retirement could be sensitive to the definition of this proxy variable. The effect of the definition of the proxy

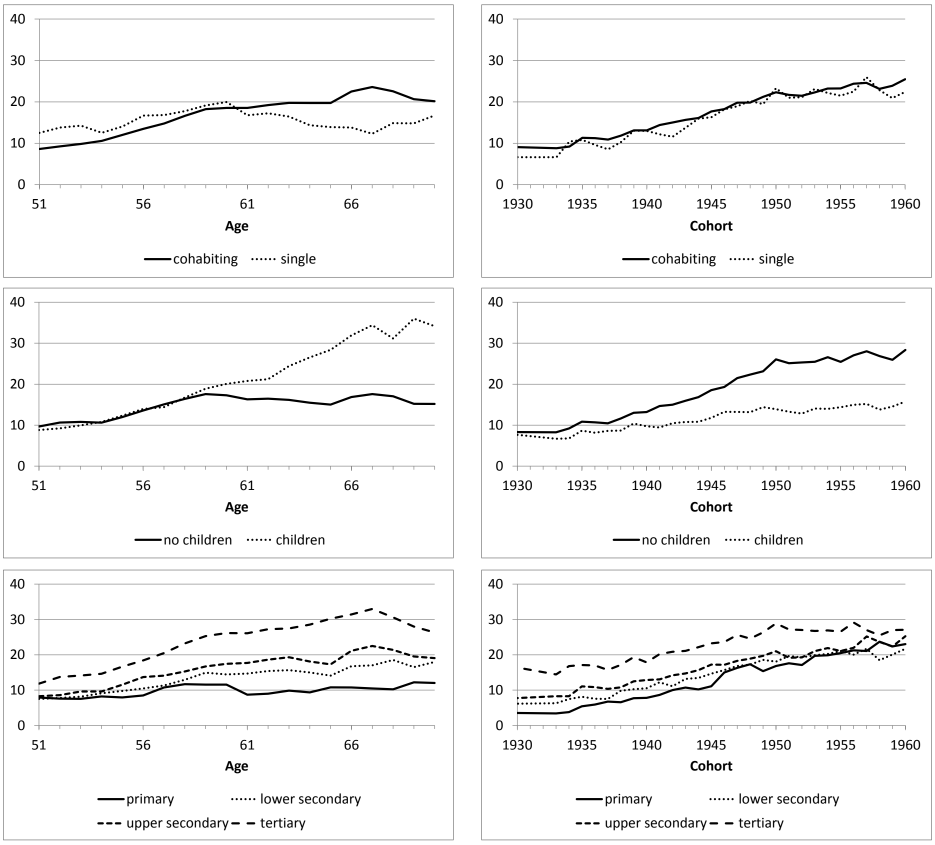


Figure 4.4: Probability of gradual retirement, by age (left panels) and cohort (right panels) in %-points under the proxy variable approach, as a function of a benchmark person. For the benchmark person the mean of the (subset of the) data is taken, except for the age (cohort) dummies. Each graph is based on two (four) separate MNL estimations, estimated on subsets of the data by: position in the household (upper panels), children (middle panels), and educational attainment (lower panels).

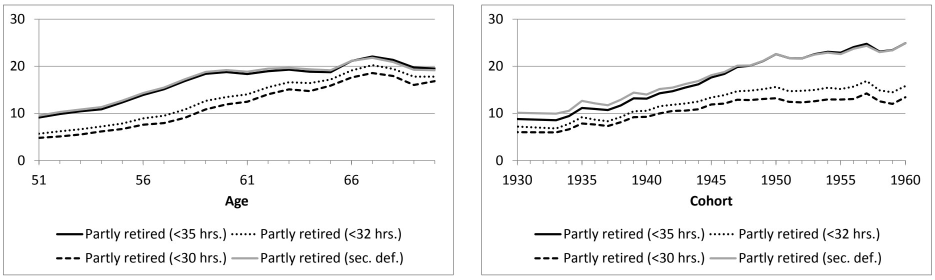


Figure 4.5: Probability of gradual retirement under different bounds on the definition of gradual retirement and a different definition of gradual retirement, by age (left panel) and cohort (right panel) in %-points under the proxy variable approach, as a function of a benchmark person. For the benchmark person the mean of the data is taken, except for the age (cohort) dummies.

variable is investigated using two different approaches. First, the cut-off point of the number of hours an individual performs paid work is adjusted to strictly less than 32 hours, and strictly less than 30 hours. In addition, the definition of gradual retirement is adjusted to:

Gradual retirement: alternative definition in this study

An individual is defined as working part-time, and thus gradual retired, if he either works strictly less than 35 hours per week, or indicates that he would like to work strictly less than 35 hours per week conditional on the event that he is currently unemployed. Notice that under this definition of gradual retirement, individuals who are unemployed, but willing to work at least one hour per week, are included in the analysis.

Graphically, the adjusted definition of a gradually retired individual hardly seems to affect the age and cohort effects (see Figure 4.5). Figure 4.5 shows, in addition, that changing the cut-off point of the number of hours worked shifts the probability of gradual retirement downwards, while hardly affecting the overall development. The downward adjustment on the probability to gradually retire when downwards adjusting the cut-off point is as expected. Fewer individuals are classified as gradually retired when the cut-off point of the number of hours worked is downward adjusted, while more individuals are classified as fully working (the number of fully retired individuals remains unaffected).

4.2 SMNL: fully working, gradually retired by self-employment or as employee, or fully retired

Labour agreements in the Netherlands make it difficult to gradually retire by phased retirement or to continue working, after the age when an individual is first eligible to an Old Age pension. Alternatively an individual could gradually retire through self-employment, as suggested by, e.g., Bloemen et al. (2014). This section investigates if there is a difference in the probability of gradual retirement through self-employment, or gradual retirement as an employee, at different ages and for different cohorts.

4.2.1 Empirical strategy

To incorporate the option to gradually retire through self-employment and to gradually retire at an employer, ideally a NL model is estimated. However, due to a lack of variables in the DLFS that differ for each alternative, estimating a NL model is infeasible. Alternatively, a MNL model with four choices, or a sequential MNL (SMNL) model could be estimated[2]. In the setting of this study, the SMNL[3] splits the decision process of gradual retirement through self-employment or at an employer into two stages (see Figure 4.6 for a schematic overview). First, the individual chooses between fully working, fully retired, or gradually retired. Next, conditional on the event that in Stage 1 the individual chose gradual retirement, he then chooses between gradual retirement through self-employment, or as an employee.

To formalize, at a given time there are J alternatives, indexed $j = 1, \dots, J$ divided into H sub choice sets, $S_1, \dots, S_H$. Assume that an individual chooses first between one of these sub choice sets (Stage 1) for some h, and afterwards chooses between alternative $j \in S_h$. In addition, assume that the decision in these stages is made according to a MNL model. Then, following Nagakura and Kobayashi (2009),

$$\text{(Stage 1)} \qquad P(y \in S_h) = \frac{\exp(\mu_i'\theta_h)}{\sum_{u=1}^{H} \exp(\mu_i'\theta_u)} \tag{4.1}$$

and

$$\text{(Stage 2)} \qquad P(y = j|S_h) = \frac{\exp(\mu_i'\kappa_j)}{\sum_{k \in S_h} \exp(\mu_i'\kappa_k)}$$

where $\mu_i(x_{it}, a_{it}, t, c_i)$ is a matrix of observables for individual i, and θ_j and κ_j are vectors that contains the parameters. Identification is achieved

[2] The MNL model and SMNL model are extreme cases of the NL model (see Train (2009) p. 80-81)

[3] For a more general discussion corresponding SMNL see, e.g., Van Ophem and Schram (1997). Notice that the elimination by aspects model by Tversky (1972) essentially is a sequential logit model.

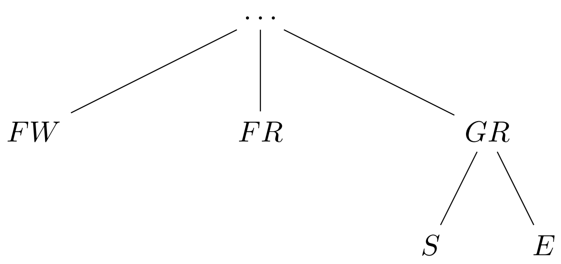


Figure 4.6: Schematic view of the choice an individual is assumed to face. FW is an abbreviation for fully working, FR for fully retired, GR for gradually retired, S for self-employed and E to indicate that an individual is gradually retired as an employee.

by normalization of $\theta_H = 0$ and $\kappa_{j_h} = 0$, where j_h is the first element in S_h. Notice that from an utility maximization perspective, contrary to a NL model, at the first stage the utility for each choice set is maximized and afterwards for each alternative at the second stage. Consequently, rather than to estimate the whole model at once, the model could be estimated in two stages (Stage 1 and Stage 2).

In this section $H = 3$ and $J = 4$, where $S_1 = \{1\}$ and $S_2 = \{2\}$. The probabilities in equation (4.1) then simplify to

$$\text{(Stage 1)} \qquad P(y = j) = \frac{\exp(\mu_i' \theta_h)}{\sum_{u=1}^{H} \exp(\mu_i' \theta_u)} \qquad \text{for } j = 1, 2.$$

and the probability of gradual retirement through self-employment ($y = 3$) or at an employee ($y = 4$) can be written as:

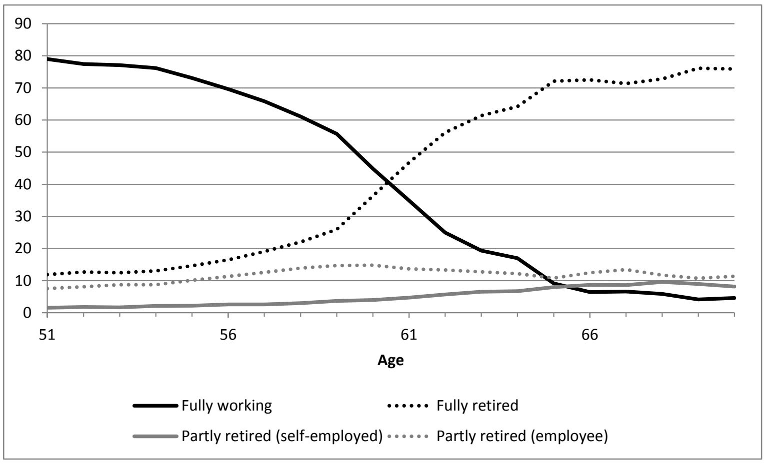
$$P(y = 3) = \frac{1}{1 + \exp(\mu_i' \kappa)} \frac{1}{1 + \exp(\mu_i' \theta_1) + \exp(\mu_i' \theta_2)}$$

$$P(y = 4) = \frac{\exp(\mu_i' \kappa)}{1 + \exp(\mu_i' \kappa)} \frac{1}{1 + \exp(\mu_i' \theta_1) + \exp(\mu_i' \theta_2)}$$

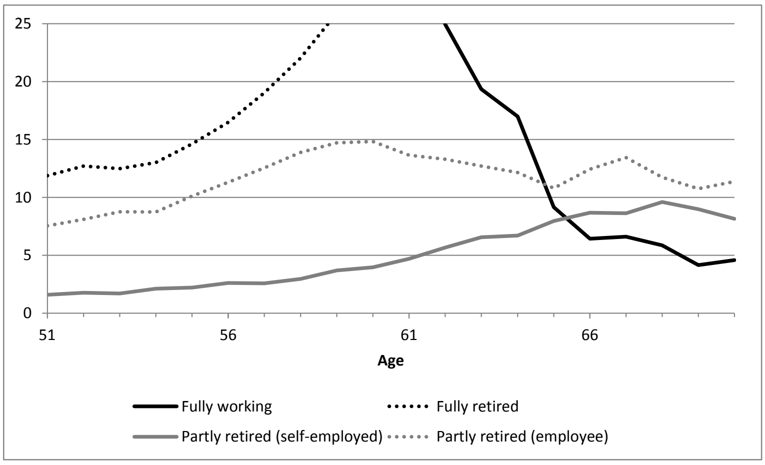
4.2.2 Results

The estimates of the previously described procedure[4] can be found in Table A.7 and Table A.8 of Appendix A. The unobserved age effects and unobserved cohort effects on the choice of gradual retirement through self-employment or as an employee are presented in Figure 4.7, respectively Figure 4.8.

[4] Also a MNL model with four choices (fully working, full retirement, gradual retirement through self-employment and gradual retirement at an employer) is estimated. Although the log likelihood (-539296.54) indicates that the MNL model fits the data better, test of the IIA property to hold, suggest that the IIA is unlikely to hold.



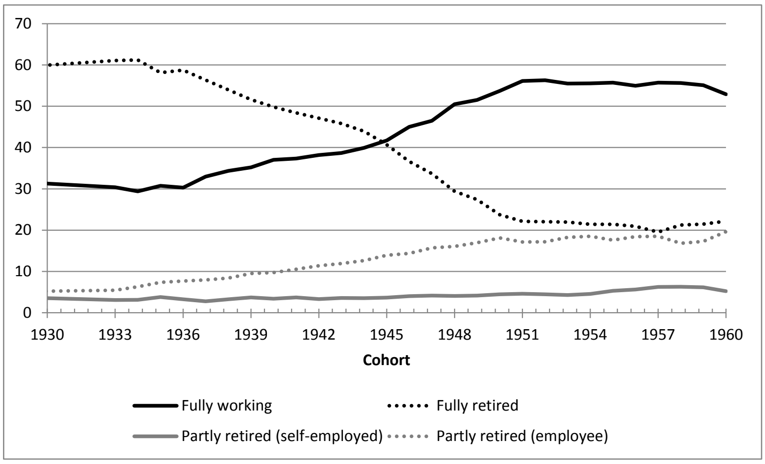
(a)



(b)

Figure 4.7: (a) Probability of full retirement, partly retired (through self-employment), partly retired (as an employee), or fully working by age (in %) under the proxy variable approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the age dummies. (b) As for (a) but for a different scaling of the y-axis.

The (unobserved) age effects of partial retirement at an employer (see Figure 4.7) peaks at the age of 60, to decrease until the age of 65, at which the probability of the choice of gradual retirement at an employer starts to increase again, until the age of 67. The probability of the choice on gradual retirement through self-employment behaves differently. Starting from the age of 57, the probability of the choice on gradual retirement through self-employment roughly linearly increases until the age of 68. Thus, while the probability of gradual retirement at an employer decreases during the ages



(a)

(b)

Figure 4.8: (a) Probability of full retirement, partly retired (through self-employment), partly retired (as an employee), or fully working by age (in %) under the proxy variable approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the cohort dummies (which include the VPL dummy). (b) As for (a) but for a different scaling of the y-axis.

starting from 60 until 66, the probability of gradual retirement through self-employment increases.

The cohort effects on the choice of gradual retirement through self-employment remain roughly constant for the cohorts starting from the cohort of 1930 until the cohort of 1952. From the cohort of 1953 onwards the choice of gradual retirement through self-employment increases until the cohort of 1957 and remains roughly constant afterwards. Furthermore, the movement of the cohort effects on the probability of gradual retirement at an employer (Figure 4.8) is roughly similar to the movement of the cohort effects on the probability of gradual retirement of Figure 4.2.

This section can only partially explain the second increase in the probability of gradual retirement as indicated by Figure 4.1 (see Section 4.1). A substantial part of the second increase can be explained by the increase in probability of gradual retirement through self-employment, the other part is likely to be induced by individuals who "unretire" (see Henkens, Van Solinge, and Van Dalen (2013) p. 54). However, the data of the DLFS does not provide a possibility to further investigate this conjecture.

4.3 MNL: gradual retirement by working a certain number of hours

This section presents the estimation results of the APC analysis by a MNL model where an individual faces four possible retirement options. Compared to the model of Section 4.1 the option of gradual retirement is divided into three categories: gradually retired by working 1-12 hours, working 13-24 hour, and working 25-34 hours. The marginal effects, the change in probability of gradual retirement (or the change in probability to fully retire) due to a change in a certain exogenous variable under the proxy variable approach are reported in Table 4.3. The effects of partly retired and fully retired add up to the marginal effects on the probability of fully working (with a plus or minus sign). In addition, a decomposition of the aggregate growth of gradual retirement in the period 1992-2012 is performed.

4.3.1 Results

As the estimated marginal effects of full retirement of Table 4.3 and Table 4.1 hardly differ, these estimates will not be discussed in more detail in this section. The probability of gradual retirement by working 1-12 hours decreases being single while the effect of being single increases the probability of gradual retirement by working 13-24 hours and gradual retirement by working 25-34 hours. The effect on the probability of gradual retirement of having children living at home is different for gradual retirement by working 1-12 hours, 13-24 hours, or 25-34 hours. Furthermore, educational attainment, when significantly different from zero at conventional levels, increases the probability of gradual retirement.

The combined probabilities of partial retirement of Figure 4.9 develops similar to the probability of partial retirement without gradual retirement split into three categories (see Figure 4.1). Separately, the probability of gradual retirement by working 1-12 hours increases starting from the age of 57 and peaks at the age of 67, while remaining roughly at the same level

Table 4.3: Marginal effects for the probability of full retirement, partly retired (1-12 hours), partly retired (13-24 hours), or partly retired (25-34 hours) (in %-points), men aged 50-70 under the proxy variable approach.[a,b,c,d]

	Fully retired		Partly retired (1-12 hours)		Partly retired (13-24 hours)		Partly retired (25-34 hours)	
	Marginal effect	t-value	Marginal effect	t-value	Marginal effect	t-value	Marginal effect	t-value
Position in household								
Cohabiting								
Single	7.73	60.74	−0.38	−6.02	0.64	8.83	1.57	18.92
Other	6.28	15.39	−0.99	−4.01	1.12	4.96	2.19	8.23
Children								
No children								
Only children < 18	−4.08	−19.57	0.05	0.48	1.12	12.47	1.33	13.64
Both children < 18 and >=18	−6.92	−25.05	−0.26	−1.90	0.39	3.04	−0.45	−3.51
Only children >= 18	−5.19	−41.83	−0.23	−3.76	−0.13	−1.90	−0.93	−11.88
Education								
Primary								
Lower secondary	−7.56	−36.56	−0.02	−0.15	−0.17	−1.33	0.67	3.99
Higher secondary	−10.07	−41.51	0.34	2.82	0.11	0.72	1.39	7.22
Tertiary	−16.10	−60.35	1.07	8.09	0.96	5.93	4.24	20.15
Age								
Dummy variables	Yes	***	Yes	***	Yes	***	Yes	***
Period (year)								
U_t	−0.05	−1.30	−0.04	−2.22	−0.04	−1.78	−0.03	−1.19
Cohort (year of birth)								
Dummy variables	Yes	***	Yes	***	Yes	***	Yes	***
VPL	−20.09	−20.53	3.50	9.35	3.14	6.09	5.24	6.46

[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

[b] The model for the probability of fully working or partly retired is estimated by a multinomial logit model with fully working as the reference category (see Table A.1 for the estimates of this model).

[c] The age effects are presented in Figure 4.1, the cohort effects in Figure 4.2 and the corresponding marginal effects can be found in Table A.3.

[d] The reported point estimates are average marginal effects. That is, for each individual the marginal effect is calculated and averaged afterwards.

Table 4.4: Probability of full retirement, partly retired (1-12 hours), partly retired (13-24 hours), partly retired (25-34 hours) or fully working (in %-points).[a]

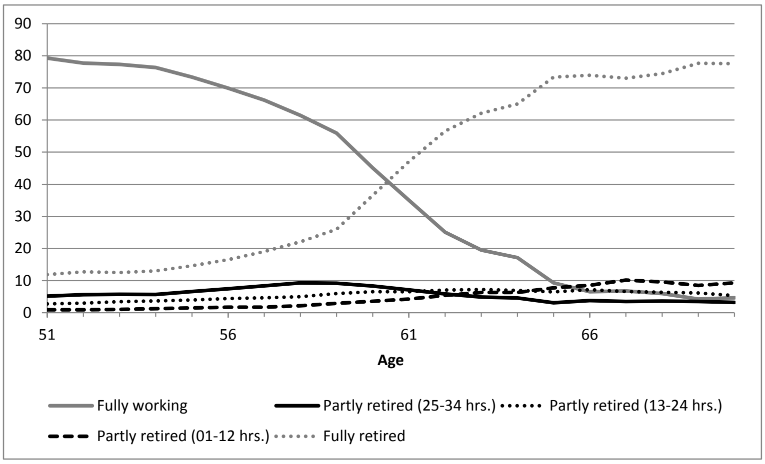
	Fully retired	Partly retired (1-12 hours)	Partly retired (13-24 hours)	Partly retired (25-34 hours)	Fully working
Age					
55	10.1	1.4	3.7	7.8	77.0
60	27.4	3.6	6.7	10.7	51.5
65	65.3	9.4	7.9	4.7	12.6
70	70.4	11.5	6.6	5.0	6.5

[a] The probabilities refer to person A with the following characteristics: year of birth 1950, higher secondary education, cohabiting, no children living at home, aggregate unemployment rate 4.75 percent.

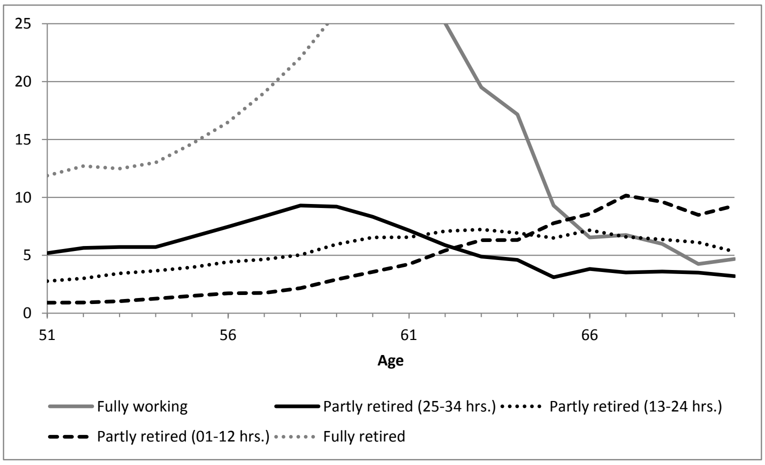
before the age of 57. The probability of gradual retirement by working 13-24 hours roughly increases until the age of 62 to remain constant afterwards. At the age of 59 the probability to gradually retire by working 25-34 hours peeks. After the age of 59 this probability decreases until the age of 66 where it roughly remains constant. This suggest that men in the Netherlands prefer, when possible, to decrease the number of hours worked per week towards the age when first eligible to receive Old Age pension benefits.

Also the combined probabilities of gradual retirement of Figure 4.10 move in correspondence with the probability of gradual retirement of Figure 4.2. The probability of gradual retirement by working 1-12 hours roughly increases linear starting from individuals born in 1930 until the cohort of 1960. Until the cohort of 1948, the probability of gradual retirement by working 13-24 hours roughly increases linearly. Starting from individuals born in 1949 the probability of gradual retirement by working 13-24 hours is roughly constant. The increase in the probability of gradual retirement by working 25-34 hours was most severe for individuals born in 1933 until the cohort of 1950. After the cohort of 1950 this probability still increases approximately linearly until the individuals born in 1960. The most likely explanation for the kink in the probability of gradual retirement by working 25-34 hours is similar to that of Section 4.1.1, namely the effect of changes in legislation which affects individuals born on or after January 1, 1950.

Figure 4.9 and Figure 4.10 only present the development of the probability of gradual retirement for the benchmark person. However, Section 4.1.3 already pointed out that the probability of gradual retirement could partially depend on the observed individual characteristics. To investigate whether the development of the probability of gradual retirement of Figure 4.9 and Figure 4.10 are roughly similar to that of a typical individual in the DLFS, a typical person ('Person A') is constructed. The effects of the birth year (cohort) and age for 'Person A' are presented in Figure 4.11 and Table 4.4. Figure 4.11 indicates that, as suggested by Section 4.1.3, the probability of gradual retirement varies with educational attainment. Furthermore, the probability of gradual retirement varies, over age (Table 4.4). Similar to



(a)

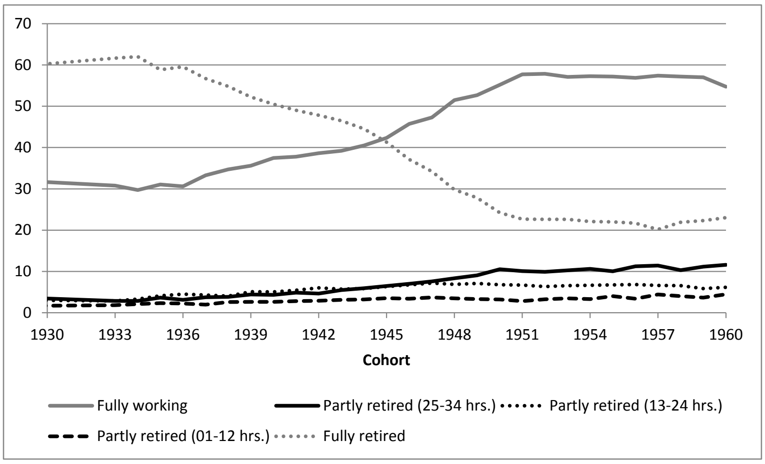


(b)

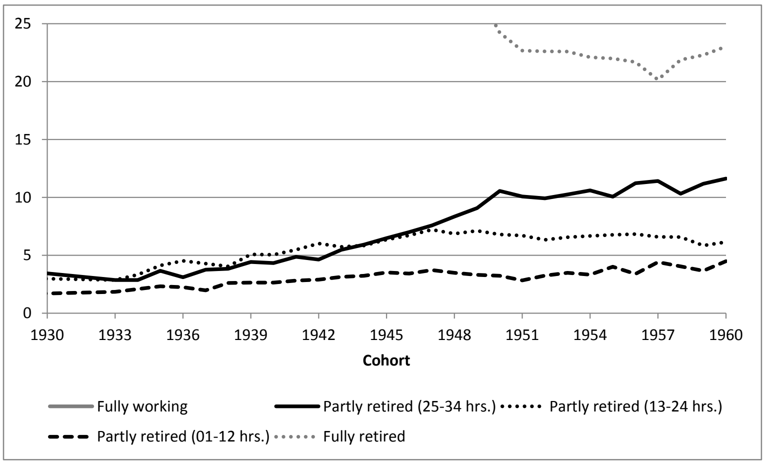
Figure 4.9: (a) Probability of full retirement, partly retired (1-12 hours), partly retired (13-24 hours), partly retired (25-34 hours), or fully working by age (in %) under the proxy variable approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the age dummies. (b) As for (a) but for a different scaling of the y-axis.

Figure 4.9 the probability to work full time decreases from the age 55 to 70, the probability of gradual retirement by working 25-34 hours peaks around the age of 60, the probability of gradual retirement by working 13-24 hours peaks around the age of 65 while the probability of gradual retirement by working 1-12 hours increases gradually until the age of 70.

To determine the contribution of the observed variables in the aggregated growth of gradual retirement in the period 1992-2012, a decomposition of gradual retirement by working 1-12 hours, gradual retirement by working



(a)



(b)

Figure 4.10: (a) Probability of full retirement, partly retired (1-12 hours), partly retired (13-24 hours), partly retired (25-34 hours), or fully working by cohort (in %) under the proxy variable approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the cohort dummies (which include the VPL dummy). (b) As for (a) but for a different scaling of the y-axis.

13-24 hours and gradual retirement by working 25-34 hours is performed. A decomposition of the aggregate growth is often performed on the basis of the marginal effects (see, e.g, Bosch et al. (2010), and Euwals, Knoef, and Van Vuuren (2011)). However, a decomposition on the basis of marginal effects is essentially a linear approximation, whereas the Multinomial Logit is non-linear. Large changes in the observed variables could therefore imply large changes in the second and higher order effects. To obtain a more

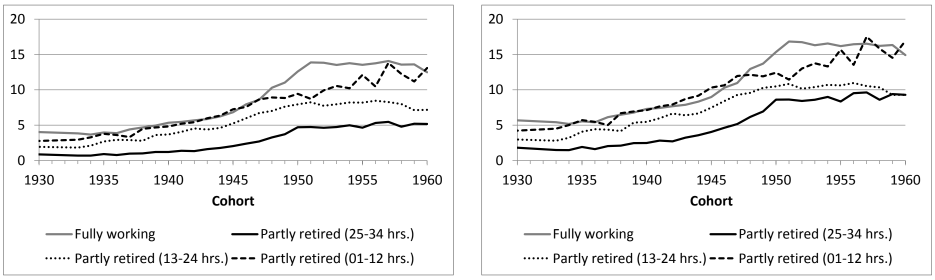


Figure 4.11: ('Left panel') Probability of partly retired (1-12 hours), partly retired (13-24 hours), partly retired (25-34 hours), or fully working by cohort (in %) under the proxy variable approach as a function of a person A. Person A has the following characteristics: age 65, higher secondary education, cohabiting, no children living at home, aggregate unemployment rate 4.75 percent. ('Right panel') As for ('Left panel') but for person B. Person B has the same characteristics as person A, but has tertiary education.

precise decomposition, I prefer the following strategy[5] over the approach on the basis of marginal effects, to determine the contribution of each of the observables.

Using the estimates, $\hat{\theta} = [\hat{\theta}_1; \hat{\theta}_2; \ldots; \hat{\theta}_5]$ of the MNL model of the current section (see Table A.9 and Table A.10) the contribution, C_j, to the change in the probability to choose option j by a group of observables z is defined by

$$C_j(\mu_{a,b}, \hat{\theta}) = \frac{\exp(\mu'_{a,b}\hat{\theta}_j)}{1 + \sum_{k=2}^{5} \exp(\mu'_{a,b}\hat{\theta}_k)} - \frac{\exp(\mu'_{a,a}\hat{\theta}_j)}{1 + \sum_{k=2}^{5} \exp(\mu'_{a,a}\hat{\theta}_k)} \quad \text{for } a = 1992; b = 2012 \qquad (4.2)$$

where $\mu_{a,b}$ is a vector that contains the vectors x_b and z_a. Vector x_b contains all explanatory variables except those in z_a, $\hat{\theta}$ is a vector of coefficients corresponding with $\mu_{a,b}$. The vectors x_b and z_a both contain the average values for one particular year (indicated by the subscript). Notice that equation (4.2) is simply the difference between the estimated probabilities of choice j under a MNL model using different combinations of observables. The results of this procedure are presented in Table 4.5.

The results of the decomposition suggest that the most important factor in explaining the growth in gradual retirement (and fully working) is the change in legislation (as described in Section 2.1.3). Besides this effect, also education and the dummy variable part of cohort effects contribute to the growth. Other effects, such as period effects, household position and children living at home, only play a minor role.

[5]Also a decomposition on the basis of the marginal effects is performed. This decomposition is, as expected, imprecise (most likely due to second and higher order effects).

Table 4.5: Decomposition of growth in the probability of full retirement, partly retired (1-12 hours), partly retired (13-24 hours), partly retired (25-34 hours) or fully working (in %-points), 1992-2012.[a]

	Fully retired	Partly retired (1-12 hours)	Partly retired (13-24 hours)	Partly retired (25-34 hours)	Fully working
Total growth 1992-2012	−45.1	1.2	2.8	7.1	34.0
Position in household	0.1	0.0	0.0	0.0	−0.1
Children	−0.8	0.0	0.1	0.1	0.6
Education	−4.7	0.1	0.2	0.7	3.7
Age					
Dummy variables	−9.4	−0.1	−0.1	0.4	9.2
Period (year)					
U_t	−0.1	0.0	0.0	0.0	0.1
Cohort (year of birth)					
Dummy variables	−4.6	−0.1	0.1	0.7	3.9
VPL	−25.5	1.4	2.5	5.1	16.5

[a] The differences between the total growth and the decomposition of the growth are spread out over all components according to the relative share of a component in total growth.

4.3.2 Sensitivity of the cut-off point

Similar to the first section of this chapter, a sensitivity analysis on the cut-off point of the definition of gradual retirement is conducted[6]. The resulting unobserved age and cohort effects are presented in Figure 4.12. Changing the cut-off point of the number of hours worked shifts the probability of partial partial retirement (25-31 hours) downwards for both the cohort and age effects. The cohort effects of partial retirement (25-31 hours) now moves roughly similar to the cohort effects of partial retirement (1-12 hours). Thus, a substantial part of men in the Netherlands is working between 32-34 hours per week.

4.4 MNP: fully working, gradually retired, or fully retired

To estimate the MNP model of Section 3.2.4, a double integral (for the bivariate normal distribution) needs to be calculated numerically. Sheppard (1900) proved that the standard bivariate normal distribution can be rewritten into an expression which only needs to evaluate a single integral. The adapted formula of a standard bivariate normal distribution forms the basis of an algorithm proposed by Drezner and Wesolowsky (1990) to approximate the cumulative distribution function of a bivariate normal distribution. This method is implemented in Matlab by Genz (2001) and is used in this study to maximize the log likelihood of equation (3.5).

[6] Notice that the (unobserved) age and cohort effects of partial retirement (13-24 hours), partial retirement (1-12 hours) and full retirement are thus unaffected.

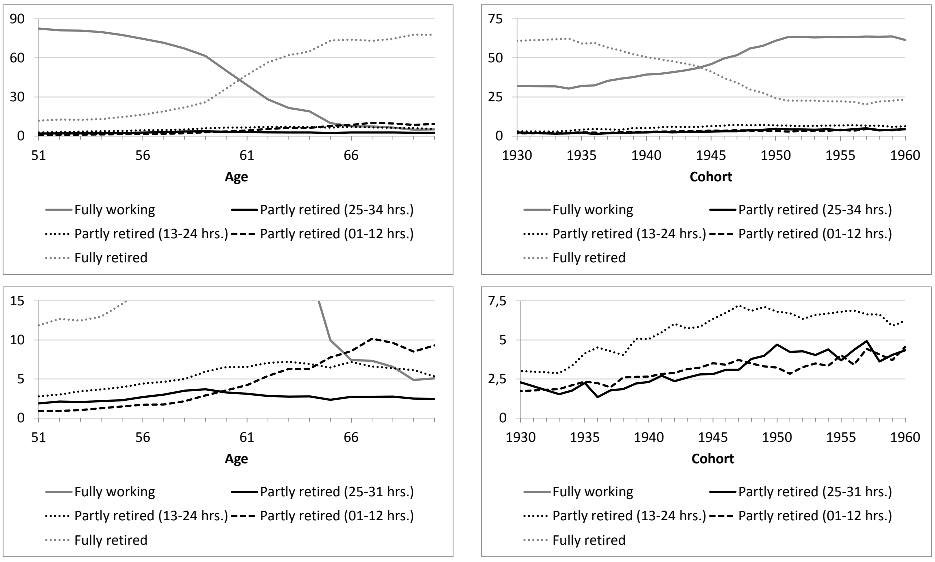


Figure 4.12: ('Upper panels') Probability of full retirement, partly retired (1-12 hours), partly retired (13-24 hours), partly retired (25-34 hours), or fully working under a strictly less than 32 hours bound on the definition of gradual retirement, by age (left panels) and cohort (right panels) in %-points under the proxy variable approach, as a function of a benchmark person. For the benchmark person the mean of the data is taken, except for the age (cohort) dummies. ('Lower panels') As for ('Upper panels') but for a different scaling of the y-axis.

The starting values of θ_2 and θ_3 for the optimization procedure are obtained by estimating a MNP model with independent identically normally distributed error terms. Hereafter, the MNP model of Section 3.2.4 is estimated under different starting values of the the correlation coefficient[7] between ϵ_2 and ϵ_3, ρ, and σ_{33}. The estimation results for different starting values of ρ and σ_{33} are presented in Table 4.6 and Table A.13. Although most of the estimates do not seem sensitive to the starting values of ρ and σ, the standard errors of the estimates are. For example, the standard error of 'Partly retired: *Position in household* Single' under Model (1) is almost 15 times smaller than the standard error of 'Partly retired: *Position in household* Single' under Model (3). Also the standard error of σ_{33} under Model (3) is roughly 20 times as large as the standard error of σ_{33} under Model (2).

The sensitivity of the standard errors of the estimates to the starting value is not completely surprising. Keane (1992) showed, by Monte Carlo simulations, that without alternative specific variables the (standard errors of the) estimates of a MNP model are likely to be sensitive to starting values of the optimization procedure. He indicates that this sensitivity "arises

[7]Recall that $\text{corr}(\epsilon_2, \epsilon_3) = \rho = \sigma_{23}/\sqrt{\sigma_{33}}$.

Table 4.6: Estimates of the multinomial probit model with fully working as the reference category, men aged 50-70 under the proxy variable approach and under different starting values of ρ and σ_{33}.[a,b]

	Model (1)		Model (2)		Model (3)	
	Estimate	std. err.	Estimate	std. err.	Estimate	std. err.
Partly retired						
Position in household						
Cohabiting						
Single	0.214	0.000700	0.213	0.001044	0.214	0.010356
Other	0.221	0.003821	0.221	0.001634	0.221	0.014255
Children						
No children						
Only children < 18	0.070	0.000998	0.070	0.000908	0.070	0.008406
Both children < 18 and >=18	−0.147	0.001210	−0.147	0.000645	−0.147	0.008839
Only children >= 18	−0.163	0.001234	−0.163	0.000917	−0.163	0.006854
Education						
Primary						
Lower secondary	−0.097	0.000723	−0.097	0.000429	−0.097	0.007700
Higher secondary	−0.061	0.000820	−0.061	0.000857	−0.062	0.007174
Tertiary	0.090	0.001215	0.090	0.000412	0.089	0.008651
Age						
Dummy variables	Yes	***	Yes	***	Yes	***
Period (year)						
U_t	−0.006	0.000710	−0.006	0.000466	−0.006	0.013670
Cohort (year of birth)						
Dummy variables	Yes	***	Yes	***	Yes	***
VPL	0.376	0.002614	0.377	0.001190	0.374	0.012073
β_0	−1.543	0.000571	−1.544	0.000402	−1.539	0.011811
Fully retired						
Position in household						
Cohabiting						
Single	0.372	0.001687	0.374	0.001156	0.369	0.013200
Other	0.300	0.002917	0.301	0.001153	0.298	0.009151
Children						
No children						
Only children < 18	−0.137	0.001847	−0.138	0.000780	−0.135	0.010462
Both children < 18 and >=18	−0.299	0.002142	−0.301	0.000612	−0.296	0.007203
Only children >= 18	−0.257	0.000824	−0.258	0.001243	−0.254	0.008428
Education						
Primary						
Lower secondary	−0.340	0.000970	−0.341	0.001457	−0.336	0.007903
Higher secondary	−0.433	0.001954	−0.434	0.000465	−0.428	0.007795
Tertiary	−0.638	0.001261	−0.640	0.000597	−0.629	0.010718
Age						
Dummy variables	Yes	***	Yes	***	Yes	***
Period (year)						
U_t	−0.004	0.000311	−0.004	0.000545	−0.004	0.002331
Cohort (year of birth)						
Dummy variables	Yes	***	Yes	***	Yes	***
VPL	−0.726	0.002023	−0.730	0.001695	−0.716	0.012364
β_0	−0.106	0.001752	−0.106	0.001196	−0.105	0.009716
ρ	0.279	0.000560	0.277	0.000912	0.287	0.005267
σ_{33}	0.834	0.000303	0.840	0.000194	0.815	0.003970
Log-likelihood	-490357.87		-490357.88		-490357.87	

[a] The starting values for Model (1) are: $\rho = 0$ and $\sigma_{33} = 1.00$, Model (2): $\rho = 0.60$ and $\sigma_{33} = 1.50$ and Model (3): $\rho = -0.80$, and $\sigma_{33} = 0.25$.

[b] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

Figure 4.13: Ratio of the probability of fully working, partly retired or fully retired between the MNL and MNP model, as a function of a benchmark person. For the benchmark person the mean of the data is taken, except for the cohort (age) dummies.

because it is difficult to disentangle the covariance parameters from regressor coefficients in such models."[8] Nevertheless, these results do provide an opportunity to inspect whether the estimated probabilities of gradual retirement by the MNP model are different from the estimated probabilities of gradual retirement by a MNL model as the estimates of Table 4.6 and Table A.13 are approximately equal.

Figure 4.13 consists of the ratio[9] of the unobserved cohort effects ('Left panel' and the ratio of unobserved age effects ('Right panel')[10]. Graphically, the unobserved cohort effects derived under the MNP model hardly differ from the unobserved cohort effects derived under the MNL model as the ratios are roughly equal to one. Also the unobserved age effects derived under the MNP model hardly differ from the unobserved age effects under the MNL model, except for the the probability of fully working after the age of 65. The difference between this ratio and 1 can graphically be perceived as large. The probability of fully working after the age of 65 under the MNP estimates is on average 6.3% after the age of 65 while under the MNL estimates the probability of fully working after the age of 65 is 5.6%.

[8]The reported standard errors under the MNP model are the square root of the diagonal of the inverse negative Hessian. The negative Hessian is an output option when using the the 'fminunc' function of the Optimization Toolbox in Matlab of which the objective can be found in Appendix D. The negative Hessian is obtained using the 'default' settings, therefore, the Hessian can be inaccurate (see MathWorks (n.d.)) which is likely to be the case for Model (1) and Model (2).

[9]To emphasize the differences between the MNP and MNL estimates, I prefer to present the ratio rather than, for example, the absolute values of the differences.

[10]Figure 4.13 is obtained in two steps. First the probabilities of fully working, partly retired or fully retired by cohort (age) as a function of the benchmark person are calculated using the estimates of the MNP (Model (3)). For the benchmark person, the mean of the whole dataset is taken, except for the cohort (age) dummies. Next, the probabilities of fully working, partly retired or fully retired by cohort (age) under the MNL model are divided by the probabilities of fully working, partly retired or fully retired under the MNP model.

This suggests that, at least for the estimates of gradual retirement, the IIA assumption may not cause any major problems in this study[11].

4.5 Sensitivity of (possible) misclassification

Although men, in contrast to women, use part-time employment as a step from fully working to full retirement (see, e.g., Been and Van Vliet (2014)), the analysis up to this point did not distinguish between men who are gradually retired, or who simply works part-time[12]. Inspection of Figure 3.1 of individuals aged 40-50 suggests a roughly constant base level of observed gradually retired individuals at these ages. A priori, one might expect that for these ages the percentage of gradually retired individuals is close to zero. Therefore, a roughly constant base level of observed gradually retired individuals might be misclassified. This section tries to assess the effects of misclassification of individuals who work part-time and are not gradually retired.

4.5.1 Empirical strategy

In order to incorporate (possible) misclassification, the multinomial logit model of Section 3.2.1 is extended to a more general model that allows for misclassification errors. The true category is denoted by $\tilde{y}_i$, and the observed response remains denoted by y_i. The link between y_i and $\tilde{y}_i$ is mostly modelled in the literature by generalizing existing models by allowing for misclassification. Following Poterba and Summers (1995) and Hausman, Abrevaya, and Scott-Morton (1998), it is initially assumed that the probability of misclassification $\alpha_{j,k} = P(y_i = j|\tilde{y}_i = k)$ only depends on k and j, but is otherwise independent of the individuals characteristics. Moreover, a similar analysis is conducted where a predefined density function is assumed to hold for the misclassification probabilities. There are in total six independent misclassification probabilities to be estimated if $J = 3$, because of an adding up condition: $\alpha_{1,j} + \alpha_{2,j} + \alpha_{3,j} = 1$ for $j = 1, \ldots, J$ where $\alpha_{j,j}$ is the probability of correct classification.

First some additional notation is introduced before specifying the probability to observe an outcome j and consequently present the log-likelihood function for observed outcomes. Denote the "true" probability of an individual i choosing alternative j in correspondence with equation (3.2) by:

[11] The strategy to obtain Figure 4.13 is also conducted for Person A of Figure 4.11. Again the cohort effects under the MNL hardly differ from the cohort effects under MNP. In correspondence with the unobserved age effects, the probability of fully working after the age of 65 seems sensitive to the model specification.

[12] Notice that the definition of gradual retirement can not distinguish between these two (sub)options.

$$p_{i,j} = P(y_i = j|\mu_i(x_{it}, a_{it}, t, c_i), \theta_j) = \frac{\exp(\mu_i'\theta_j)}{1 + \sum\limits_{k=2}^{J} \exp(\mu_i'\theta_k)}$$

where the base category is fully working, thus $\theta_1 = 0$. The vector of "true" probabilities is denoted by: $p_i = [p_{i,1}; \ldots; p_{i,3}]$. The matrix of the misclassification probabilities, A, is defined as:

$$A = \begin{bmatrix} 1 - \alpha_{2,1} - \alpha_{3,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & 1 - \alpha_{1,2} - \alpha_{3,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & 1 - \alpha_{1,3} - \alpha_{2,3}. \end{bmatrix}$$

The vector of observed probabilities of an individual i is simply the matrix multiplication of A and p_i:

$$\tilde{p}_i = Ap_i \tag{4.3}$$

As an example, the observed probability of an individual i being gradually retired, $j = 2$, can be written as:

$$\tilde{p}_{i,2} = \frac{\alpha_{2,1}\exp(\mu_i'\theta_1) + (1 - \alpha_{1,2} - \alpha_{3,2})\exp(\mu_i'\theta_2) + \alpha_{2,3}\exp(\mu_i'\theta_3)}{1 + \sum\limits_{k=2}^{J} \exp(\mu_i'\theta_k)}$$

This and analogous expressions for the other observed probabilities form the basis of the log-likelihood function. Let

$$y_{ij} = 1(y_i = j),$$

where $1(I)$ is the indicator function which equals one if I is true and zero otherwise. The log-likelihood function is defined by:

$$\mathscr{L}(\alpha, \theta_1, \theta_2) = \sum_{i=1}^{n} \sum_{j=1}^{3} y_{ij} \log(\tilde{p}_{ij})$$

where the total number of observations in the sample is n.

The estimation results are derived under the assumption that misclassification only can occur by individuals observed as gradually retired, but are in this model "truly" classified as (fully) working. An individual is observed as fully working when he is part of the labour force and in addition performs paid work at least 35 hours per week. As these two conditions are consistently available in the data, misclassification with respect to this category and another category is unlikely, thus $\alpha_{1,2} = 0$, $\alpha_{1,3} = 0$ and $\alpha_{3,1} = 0$. A similar argument holds to observe an individual as gradually retired, thus it is unlikely that an individual is observed as fully retired while he truly is gradually retired, or vice versa ($\alpha_{3,2} = 0$ and $\alpha_{2,3} = 0$). There is, however, ambiguity between men working part-time or are gradually retired. Therefore, I assume that A is defined as:

$$A = \begin{bmatrix} 1-\alpha_{2,1} & 0 & 0 \\ \alpha_{2,1} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.4}$$

Identification of this model, Model (m.1), is achieved by assuming, next to the conditions necessary for identification of the multinomial logit model, that $\alpha_{2,1} < 1$. The motivation for this additional (monotonicity) condition can be found in Hausman et al. (1998). The Matlab code to estimate Model (m.1) by maximum likelihood estimation can be found in Appendix D of this study.

A drawback of Model (m.1) is that the misclassification probability is independent of the age of an individual. It is more likely that an individual of age 50 is observed in the data as gradually retired while he truly prefers to works part-time over gradual retirement, compared to an individual with the age of 65. I will now present Model (m.2) that takes the age of an individual into account.

To learn about the effects of possible misclassification of men who work part-time rather than being gradually retired conditional on the age of an individual, the distribution of $\alpha_{2,1}$ is specified beforehand. Let $\alpha_{(2,1),age}$ denote the probability of misclassification at a given age defined by

$$\alpha_{(2,1),age} = \tilde{\alpha}_{(2,1)} \frac{|GR|_{age} + |FW|_{age}}{|GR|_{age}} \tag{4.5}$$

where $|\cdot|_{age}$ denotes the total number of observed individuals who are gradually retired (GR), or fully working (FW) at a given age in the dataset. The variable $\tilde{\alpha}_{(2,1)}$ denotes the percentage of working individuals who are assumed to work part-time, irrespective of the number of hours worked per week and is assumed not to change at different ages and is equal for each cohort.

For a given $\tilde{\alpha}_{(2,1)}$ and a given (estimated) $\tilde{p}_i$ it is now possible by using equation (4.5) with A defined by equation (4.4), to obtain the "true" probabilities. Under these assumption, $p_{i,1}$ can be obtained using the relationship:

$$p_{i,1} = \min(\tilde{p}_{i,1}/(1 - \alpha_{(2,1),age_i}), 1)$$

with age_i the age of individual i. Then $p_{i,2}$ is calculated by

$$p_{i,2} = \max(\tilde{p}_{i,2} - p_{i,1}\alpha_{(2,1),age_i}, 0),$$

and observe that $p_{i,3} = \tilde{p}_{i,3}$. The maximum and minimum operators are used to ensure that the "true" probabilities will not become negative or greater than 1. Finally, it should be emphasized that the results of Model (m.2) itself are not estimated, but are derived under the estimation results of a multinomial logit model with predefined misclassification probabilities.

Table 4.7: Probabilities of misclassification at different ages, $\alpha_{(2,1),age}$, for different $\tilde{\alpha}_{(2,1)}$.

	$\tilde{\alpha}_{(2,1)} = 0.04$	$\tilde{\alpha}_{(2,1)} = 0.06$	$\tilde{\alpha}_{(2,1)} = 0.08$	$\tilde{\alpha}_{(2,1)} = 0.10$
Age 50	0.369	0.554	0.738	0.923
Age 51	0.360	0.540	0.720	0.900
Age 52	0.337	0.505	0.673	0.841
Age 53	0.322	0.483	0.644	0.805
Age 54	0.310	0.465	0.620	0.775
Age 55	0.269	0.404	0.539	0.673
Age 56	0.233	0.350	0.467	0.584
Age 57	0.209	0.313	0.418	0.522
Age 58	0.179	0.269	0.359	0.449
Age 59	0.156	0.234	0.312	0.390
Age 60	0.132	0.198	0.264	0.330
Age 61	0.113	0.169	0.226	0.282
Age 62	0.090	0.136	0.181	0.226
Age 63	0.079	0.118	0.158	0.197
Age 64	0.075	0.113	0.151	0.188
Age 65	0.060	0.089	0.119	0.149
Age 66	0.053	0.079	0.106	0.132
Age 67	0.053	0.079	0.106	0.132
Age 68	0.052	0.078	0.105	0.131
Age 69	0.050	0.075	0.100	0.124
Age 70	0.051	0.077	0.103	0.128

4.5.2 Results

Under the assumption that misclassification will only occur by individuals observed as gradually retired, but are in this model "truly" classified as fully working, the estimate of $\alpha_{2,1}$ equals 0.001 and is not statistically significantly different from zero at conventional levels. The estimation results are roughly comparable to the estimation results without accounting for possible misclassification.

Investigating the data and Figure 3.1, at the ages between 40 and 50, suggest that $\tilde{\alpha}_{(2,1)}$ could be, at least for these ages, around 6%. Moreover, for $\tilde{\alpha}_{(2,1)} = 0.06$ a substantial part of individuals at the age of 50 is assumed to be misclassified (which could be realistic), while the group of individuals misclassified at the age of 70 remains relatively small. Nevertheless, other possible values of $\tilde{\alpha}_{(2,1)}$ will be investigated as well. The probabilities of $\alpha_{(2,1),age}$ are presented in Table 4.7. If $\tilde{\alpha}_{(2,1)} = 0.10$, around 95% of all gradually retired individuals are assumed to be misclassified at the age of 50 and almost 13% at the age of 70. Using these misclassification probabilities in combination with the estimation results of the multinomial logit model of Section 4.1.1 it is possible to derive "true" probabilities of gradual retirement. Figure 4.14 and Figure 4.15 present the "true" probability of

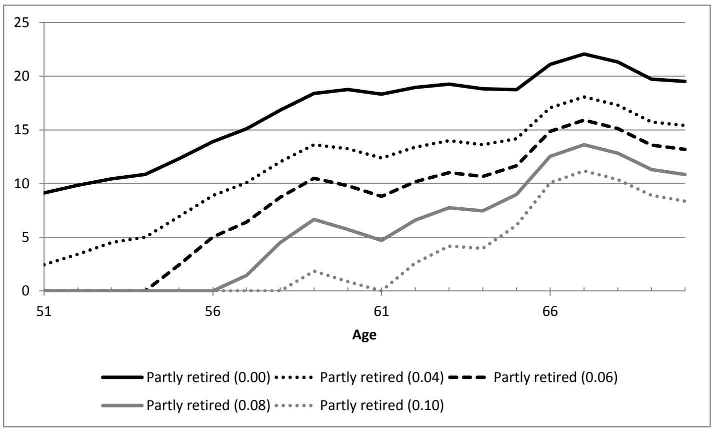


Figure 4.14: "True" probability of gradual retirement for different values of $\tilde{\alpha}_{(2,1)}$ by age (in %) under the proxy variable approach as a function of a benchmark person. For the benchmark person the mean of the whole dataset is taken, except for the age dummies.

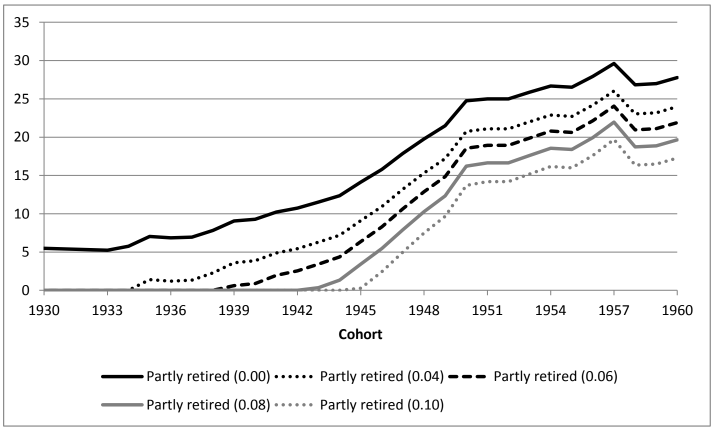


Figure 4.15: "True" probability of gradual retirement for different values of $\tilde{\alpha}_{(2,1)}$ by cohort (in %) under the proxy variable approach as a function of person A. Person A has the following characteristics: age 65, higher secondary education, cohabiting, no children living at home, aggregate unemployment rate 4.75 percent.

gradual retirement for different values of $\tilde{\alpha}_{(2,1)}$ using the earlier described methodology for Model (m.2).

These figures suggest that the probabilities without taking misclassification into account should be considered as an upper bound for the probability of gradual retirement. Although the estimated probability of gradual retirement overall decreases when taking misclassification into account, even in the more extreme scenarios the age and cohort effects on gradual retirement remain present. In addition, for the early ages of an individual, the magnitude of the effect of misclassification on the probability of gradual retirement is larger than the effect on the probability of gradual retirement at later ages. Finally, the effects for increasing $\tilde{\alpha}_{(2,1)}$ are expected beforehand. If more gradually retired individuals are assumed to be misclassified, by construction the probability of an individual to "truly" gradually retire will be low.

Chapter 5

Discussion and conclusion

This study investigated the developments of gradual retirement of men aged 50-70, using data from the DLFS 1992-2012. Descriptive statistics suggest that there is a (modest) increase in the prevalence of gradual retirement over successive cohorts, at the expense of full retirement. The change in behaviour of older men in the Netherlands to delay full retirement is in line with Van Vuuren and Deelen (2009) who concluded that the participation rate of men aged 50-69 increased over successive cohorts. To assess the probability of gradual retirement of men aged 50-70, and to gain insight in underlying trends of cohort and age effects, a multinomial logistic age-period-cohort model was estimated.

Under the assumption that an individual faces three choices (gradual retirement, full retirement, or fully working), the results indicate that for individuals born between 1933-1950 the (unobserved) cohort effects of gradual retirement roughly linearly increased. The individuals born in 1933 are the first affected by changes in legislation towards facilitating early retirement plans, while the individuals born in 1949 are the last able to utilize these plans (see Kerkhofs et al. (2009)). Hence, it is likely that the increase in the prevalence of gradual retirement can (partly) be explained by these policy changes. Also for cohorts born in, or after, the 1950s gradual retirement became more prevalent, but slower compared to cohorts born before the 1950s. The results, in addition, indicate that the (unobserved) age effects increase until the age of 59, at which they remain constant until the age of 65. After the age of 65, the (unobserved) age effects peak at the age of 67. These results need to be interpreted with care as they may pick up time trends where this study did not correct for. An example of such time trend is the increase in real wage. Using educational attainment as an explanatory variable, this study tries to take this effect (partly) into account. However, Groot and De Groot (2011) point out that, for at least the 2000-2008 period, real wages increased for most levels of education in the Netherlands.

In order to find out if the peak at the probability to gradually retire at the age of 67 could be explained by individuals who gradually retire through self-employment, a sequential multinomial logit model was estimated. The results indicate that the age effects are different for individuals who gradually retire through self-employment and individuals who gradually retire at an employer. The age effects of the latter now peak at the age of 60, to decrease until the age of 65, and to peak again at the age of 68. The age effects of gradual retirement through self-employment, however, increase from the age of 55 onwards until the age of 68. Thus, while the estimated probability of gradual retirement at an employer decreases for individuals aged 60-65, the estimated probability of gradual retirement through self-employment increases. Notice that, as transitions are not observed, it remains ambiguous whether this implies that self-employed individuals (gradually) retire later, or that composition changes induce this result. The results do suggest that, as the unobserved cohorts are roughly constant for individuals born between 1930-1955, gradual retirement through self-employment did not become more prevalent as an alternative pathway towards full retirement.

Apart from individuals who gradually retire through self-employment, also individuals who gradually retire by working 1-12 hours, gradually retire by working 13-24 hours, or gradually retire by working 25-34 hours are of interest. They could provide useful insights in the preferences of individuals from their career job towards full retirement. The results indicate, under the assumption that an individual faces five choices (thus without explicitly taking the possibility of gradual retirement through self-employment into account), that the estimated probability of gradual retirement by working 25-34 hours peaks at the age of 58, gradual retirement by working 13-24 hours peaks at the age of 62, and gradual retirement by working 1-12 hours peaks at the age of 67. Thus, when an individual becomes older, he appears to prefer, when possible, to reduce the number of hours worked. The (unobserved) cohort effects of gradual retirement by working 13-24, or by working 25-34 hours, develop similar to the cohort effects of gradual retirement by working 1-34 hours. The unobserved cohort effects of gradual retirement by working 1-12 hours per week, however, appear to be steadily increasing for cohorts born between 1930-1960. Sensitivity analysis by changing the cut-off point of the number of hours worked indicates that the estimated profile of the unobserved cohort effects of gradual retirement by working 24-31 hours per week becomes similar to the estimated profile of the unobserved cohort effects of gradual retirement by working 1-12 hours per week. This indicates that a considerable amount of individuals are working 32-34 hours. Furthermore, the results suggest that (ceteris paribus) higher educated individuals are more likely to gradually retire.

A decomposition of the aggregate growth of gradual retirement in the 1992-2012 period indicates that the most important factors in explaining the (modest) increase of gradual retirement are the unobserved cohort effects. The increase in gradual retirement by working 1-12 hours, or by working 13-

24 hours, is completely explained by the unobserved cohort effects, whereas for gradual retirement by working 25-34 hours also the increase in educational attainment is fairly important.

Although the framework of this study provides the opportunity to the estimate the profiles of both the unobserved age and cohort effects, future research is necessary to identify (components) that make up these effects. For example, the financial resources of the household where the individual lives in are likely to explain part of the unobserved age effects. Also, when relevant, the partner of the individual could be important in the timing of the decision to (gradually) retire.

The definition of gradual retirement in this study solely depends on the number of hours worked per week. Although this definition of gradual retirement provides the possibility to investigate the trends of gradual retirement, a direction for future research is to conduct an APC analysis under a definition of gradual retirement that takes transitions from the career job towards full retirement into account. This would require panel data, thereby not only allowing for more flexible model specifications, but also providing an opportunity to investigate the importance of "unretirement and to investigate whether self-employed individuals (gradually) retire later, or that compositional changes are an important factor around the age when an individual is first eligible to an Old Age pension

Gradual retirement through self-employment could be an attractive alternative for individuals who cannot, for various reason, gradually retire at an employer, from a theoretical point of view. However, this study does not find solid evidence that gradual retirement through self-employment became more prevalent over successive cohorts born between 1930-1955. Future research could investigate if for more recent cohorts this trend still holds, and which changes could be implemented to make this a more attractive alternative pathway towards full retirement. Another result of this study is the decrease in the prevalence of full retirement for individuals born between 1933-1950. This decrease of the unobserved cohort effects is in correspondence with changes in legislation. Although these changes are likely to explain (part) of this decrease, future research is required to obtain an answer to this conjecture. Acceptance of the hypothesis implies that, concerning (gradual) retirement, individuals are sensitive to the financial incentives associated with these (early) retirement plans.

To conclude, although gradual retirement became more prevalent over successive cohorts, it is unlikely that gradual retirement will become even more prevalent, unless effective policy measures are implemented or there occurs a substantial shift in social norms.

References

Been, J., & Knoef, M. (2015). *The Necessity of Self-Employment Towards Retirement: Evidence from Labor Market Dynamics and Search Requirements in Unemployment Benefits* (Discussion Paper No. 04/2015-009). Netspar.

Been, J., & Van Vliet, O. (2014). *Early Retirement Across Europe. Does Non-Standard Employment Increase Participation of Older Workers?* (Discussion Paper No. 10/2014-044). Netspar.

Bloemen, H., Hochguertel, S., & Zweerink, J. (2014). *Gradual retirement: A pathway with a future?* (Panel Paper No. 41). Netspar.

Bosch, N., Deelen, A., & Euwals, R. (2010). Is Part-time Employment Here to Stay? Working Hours of Dutch Women over Successive Generations. *Labour*, *24*(1), 35–54.

Bosworth, B., & Burtless, G. (2010). *Recessions, Wealth Destruction, and the Timing of Retirement* (Working Paper No. 2010-22). Boston College Retirement Research Center.

Brant, R. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics*, *46*(4), 1171–1178.

Brückner, H., & Mayer, K. U. (2005). De-Standardization of the Life Course: What it Might Mean? And if it Means Anything, Whether it Actually Took Place? *Advances in Life Course Research*, *9*, 27–53.

Bruinshoofd, A., & Grob, S. (2006). *Do changes in pension incentives affect retirement? A stated preferences approach to Dutch retirement consideration* (Working Paper No. 115). Netherlands Central Bank, Research Department.

Brunello, G., & Langella, M. (2013). Bridge jobs in Europe. *IZA Journal of Labor Policy*, *2*(1), 1–18.

Cahill, K., Giandrea, M., & Quinn, J. (2015). 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.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Charles, K. K., & DeCicca, P. (2007). Hours Flexibility and Retirement. *Economic Inquiry*, *45*(2), 251–267.

Cheng, S., & Long, J. S. (2007). Testing for IIA in the Multinomial Logit Model. *Sociological Methods & Research*, *35*(4), 583–600.

CPB. (2014). *Werkgelegenheidseffecten wetsvoorstel 'werken na de AOW-gerechtigde leeftijd'* (CPB-Notitie). The Hague: CPB.

Daganzo, C. (1979). *Multinomial Probit: The Theory and Its Application to Demand Forcasting*. Academic Press.

Dansie, B. (1985). Parameter Estimability in the Multinomial Probit Model. *Transportation Research B: Methodological*(25), 1–12.

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

Dorn, D., & Sousa-Poza, A. (2010). 'Voluntary' and 'involuntary' early retirement: an international analysis. *Applied Economics*, *42*(4), 427–438.

Drezner, Z., & Wesolowsky, G. (1990). On the computation of the bivariate normal integral. *Journal of Statistical Computation and Simulation*, *35*, 101–107.

Euwals, R., Knoef, M., & Van Vuuren, D. (2011). The trend in female labour force participation: what can be expected for the future? *Empirical Economics*, *40*(3), 729–753.

Euwals, R., Van Vuren, A., & Van Vuuren, D. (2011). *The Decline of Early Retirement Pathways in the Netherlands: An Empirical Analysis for the Health Care Sector* (Discussion Paper No. 06/2011-057). Netspar.

Euwals, R., van Vuuren, D., & Wolthoff, R. (2010). Early Retirement Behaviour in the Netherlands: Evidence From a Policy Reform. *De Economist*, *158*(3), 209–236.

Fienberg, S. E., & Mason, W. M. (1978). Identification and Estimation of Age-Period-Cohort Models in the Analysis of Discrete Archival Data. *Sociological Methodology*, *8*, 1–67.

Fischer, J. A., & Sousa-Poza, A. (2006). *The Institutional Determinants of Early Retirement in Europe* (Discussion Paper No. 2006-08). University of St. Gallen.

Fouarge, D., De Grip, A., & Montizaan, R. (2012). *Hoe gevoelig is de uittredeleeftijd voor veranderingen in het pensioenstelsel?* (Design Paper No. 7). Netspar.

Genz, A. (2001). *Bivariate Normal culumaltive distribution function* (Matlab code). Washington State University. Retrieved from `http://www.maths.lth.se/matstat/wafo/documentation/wafodoc/wafo/trgauss/bvnormcdf.html.`

Ghent, L. S., Allen, S. G., & Clark, R. L. (2001). The Impact of a New Phased Retirement Option on Faculty Retirement Decisions. *Research on Aging*, *23*(6), 671–693.

Giandrea, M. D., Cahill, K. E., & Quinn, J. F. (2009). Bridge Jobs A Comparison Across Cohorts. *Research on Aging*, *31*(5), 549–576.

Giandrea, M. D., Cahill, K. E., & Quinn, J. F. (2013). *New Evidence on Self-Employment Transition Among Older Americans with Career Jobs.* (Working Paper No. 463). U.S. Bureau of Labor Statistics.

Gielen, A. C. (2009). Working hours flexibility and older workers labor supply. *Oxford Economic Papers*, *61*(2), 240–274.

Gorodnichenko, Y., Song, J., & Dmitriy, S. (2013). *Macroeconomic Determinants of Retirement Timing* (Working Paper No. 19638). NBER.

Groot, S., & De Groot, H. (2011). *Wage inequality in the Netherlands: Evidence, trends and explanations* (Discussion Paper No. 186). CPB.

Hausman, J., Abrevaya, J., & Scott-Morton, F. (1998). Misclassification of the Dependent Variable in a Discrete-response Setting. *Journal of Econometrics*, *87*, 239-269.

Hausman, J., & McFadden, D. (1984). Specification Tests for the Multinomial Logit Model. *Econometrica*, *52*(5), 1219–1240.

Heckman, J., & Robb, R. (1985). Cohort Analysis in Social Research. In W. M. Mason & S. E. Fienberg (Eds.), (p. 137-150). New York: Springer-Verlag.

Henkens, K., Van Solinge, H., & Van Dalen, H. (2013). *Doorwerken over de drempel van het pensioen.* Amsterdam: Amsterdam University Press.

Hochguertel, S. (2010). *Self-Employment Around Retirement Age* (Discussion Paper No. 10-067/3). Tinbergen Institute.

Kantarci, T., & van Soest, A. (2008). Gradual retirement: preferences and limitations. *De Economist*, *156*(2), 113–144.

Kapteyn, A., Alessie, R., & Lusardi, A. (2005). Explaining the wealth holdings of different cohorts: Productivity growth and Social Security. *European Economic Review*, *49*(5), 1361–1391.

Keane, M. (1992). A Note on Identification in the Multinomial Probit Model. *Journal of Business & Economic Statistics*, *10*(2), 193–200.

Kerkhofs, M., Fouarge, D., & Ester, P. (2009). Financiële prikkels en gepreffereerde pensioenleeftijd. *Tijdschrift voor Arbeidsmarktvraagstukken*, *25*(1), 20–37.

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

Klein, R., & Spady, R. (1993). An Efficient Semiparametric Estimator for Binary Response Models. *Econometrica*, *61*(2), 387–421.

Kneib, T., Baumgartner, B., & Steiner, W. (2007). Semiparametric multinomial logit models for analysing consumer choice behaviour. *Advances in Statistical Analysis*, *91*(3), 225–224.

Lee, L. (1995). Semiparametric maximum likelihood estimation of polychotomous and sequential choice models. *Journal of Econometrics*, *65*(2), 381–428.

Li, B. (2011). The multinomial logit model revisited: A semi-parametric approach in discrete choice analysis. *Transportation Research B: Methodological*, *45*(3), 461–473.

Maestas, N. (2010). Back to Work Expectations and Realizations of Work after Retirement. *Journal of Human Resources*, *45*(3), 718–748.

Mason, K. O., Mason, W. M., Winsborough, H. H., & Poole, W. K. (1973). Some Methodological Issues in Cohort Analysis of Archival Data. *American Sociological Review*, *38*(2), 242–258.

MathWorks. (n.d.). *Hessian.* Retrieved from `http://nl.mathworks.com/help/optim/ug/hessian.html.` (Retrieved June 5, 2015.)

McFadden, D. (1973). *Frontiers in econometrics* (P. Zarembka, Ed.). Academic press.

McFadden, D. (1978). *Spatial Interaction The and Planning Models* (A. Karlquist, L. Lundquist, F. Snickars, & J. Weibull, Eds.). Amsterdam: North-Holland.

Nagakura, D., & Kobayashi, M. (2009). Testing the Sequential Logit Model Against the Nested Logit Model. *Japanese Economic Review*, *60*(3), 345–361.

O Brien, R. M. (2000). Age Period Cohort Characteristic Models. *Social Science Research*, *29*(1), 123–139.

Poterba, J. M., & Summers, L. H. (1995). Unemployment Benefits and Labor Market Transitions A Multinomial Logit Model with Errors in Classification. *The Review of Economics and Statistics*, *77*(2), 207–216.

Revelt, D., & Train, K. (1998). Mixed Logit with Repeated Choice: Households' Choice of Appliance Efficiency Level. *The Review of Economics and Statistics*, *80*(4), 647–657.

Rose, J., Hess, S., Bliemer, M., & Daly, A. (2011). *The Impact of Varying the Number of Repeated Choice Observations on the Mixed Multinomial Logit Model.* (Paper presented at the 90th annual meeting of the Transportation Research Board TRB, Washington, DC, United States, February 27, 2011)

Ruhm, C. J. (1990). Bridge Jobs and Partial Retirement. *Journal of Labor Economics*, *8*(4), 482–501.

Scott, J. (2004). *Is Phased Retirement a State of Mind? the Role of Individual Preferences in Retirement Outcomes.* (Paper presented at the Population Association of America Annual Meeting, Boston)

Sheppard, W. (1900). On the calculation of the double integral expressing normal correlation. *Transactions of the Cambridge Philosophical Society*, *19*, 23–69.

Small, K. A., & Hsiao, C. (1985). Multinomial Logit Specification Tests. *International Economic Review*, *26*(3), 619–627.

Stern, S. (1989). Rules of thumb for comparing multinomial logit and multinomial probit coefficients. *Economics Letters*, *31*(3), 235–238.

Thomson, J. (2007). *The Transition of Older Australian Workers to Full and Partial Retirement* (Research paper No. 1005). University of Melbourne.

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

Tversky, A. (1972). Elimination by Aspects: A Theory of Choice. *Psychological Review*, *79*(4), 281–299.

Van Ophem, H., & Schram, A. (1997). Sequential and Multinomial Logit: A Nested Model. *Empirical Economics*, *22*(1), 131–152.

Van Soest, A., Kapteyn, A., & Zissimopoulos, J. M. (2007). *Using stated preferences data to analyze preferences for full and partial retirement* (Discussion Paper No. 2785). IZA.

Van Soest, A., & Vonkova, H. (2014). How sensitive are retirement decisions to financial incentives? A stated preference analysis. *Journal of Applied Econometrics*, *29*(2), 246–264.

Van Vuuren, D. (2014). Flexible retirement. *Journal of Economic Surveys*, *28*(3), 573–593.

Van Vuuren, D., & Deelen, A. (2009). De participatiegroei van ouderen ontrafeld. *Economisch Statistische Berichten*, *4551*, 10-12.

Verbeek, M. (2012). *A Guide to Modern Econometrics*. John Wiley & Sons.

Vroom, R., Van Sonsbeek, J.-M., & Albas, R. (2012). Deeltijdpension in Nederland verkleint arbeidsaanbod. *Economisch Statistische Berichten*, *4645*, 612–615.

Wadensjö, E. (2006). Part-Time Pensions and Part-Time Work in Sweden. *European Papers on the New Welfare*, *6*, 29–45.

Wolfe, R., & Gould, W. (1998). An approximate likelihood-ratio test for ordinal response models. *Stata Technical Bulletin*, *7*(42).

Zhan, Y., & Wang, M. (2015). *Bridge Employment: Conceptualizations and New Directions for Future Research* (P. M. Bal, D. T. Kooij, & D. M. Rousseau, Eds.). Springer International Publishing.

Zissimopoulos, J., & Karoly, L. (2009). Labor-Force Dynamics at Older Ages: Movements Into Self-Employment for Workers and Nonworkers. *Research on Aging*, *31*(1), 89–111.

Appendix A

Extended estimation results

Table A.1: Estimates of the multinomial logit model with fully working as the reference category, men aged 50-70 under the proxy variable approach.[a]

	Partly retired		Fully retired	
	Estimate	t-value	Estimate	t-value
Position in household				
Cohabiting				
Single	0.474	40.13	0.763	70.45
Other	0.488	12.97	0.644	19.27
Children				
No children				
Only children < 18	0.144	10.39	−0.281	−17.23
Both children < 18 and >=18	−0.273	−15.20	−0.641	−30.36
Only children >= 18	−0.324	−30.67	−0.516	−51.81
Education				
Primary				
Lower secondary	−0.258	−12.13	−0.673	−38.53
Higher secondary	−0.219	−8.92	−0.851	−41.51
Tertiary	0.024	0.89	−1.229	−54.32
Age				
Dummy variables	Yes	***	Yes	***
Period (year)				
U_t	−0.018	−5.12	−0.008	−2.69
Cohort (year of birth)				
Dummy variables	Yes	***	Yes	***
VPL	0.610	5.77	−1.392	−14.48
β_0	−2.457	−22.84	−0.209	−2.24
Log-likelihood	-490443.74			

[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

Table A.2: Extended estimation results for Table A.1: estimates of the multinomial logit model for age and cohort effects.

	Partly retired		Fully retired	
	Estimate	t-value	Estimate	t-value
Age 50				
Age 51	0.050	2.23	0.082	3.07
Age 52	0.148	6.43	0.170	6.35
Age 53	0.210	9.05	0.157	5.81
Age 54	0.261	11.11	0.210	7.83
Age 55	0.429	18.27	0.368	13.92
Age 56	0.599	25.52	0.537	20.64
Age 57	0.737	31.10	0.735	28.64
Age 58	0.920	38.57	0.958	37.56
Age 59	1.102	45.69	1.123	47.88
Age 60	1.337	53.80	1.767	69.91
Age 61	1.565	60.21	2.269	88.45
Age 62	1.934	70.53	2.787	104.83
Age 63	2.204	75.16	3.131	112.15
Age 64	2.312	74.30	3.307	113.68
Age 65	2.926	78.61	4.042	117.09
Age 66	3.398	74.86	4.400	104.05
Age 67	3.415	69.77	4.357	96.46
Age 68	3.502	62.91	4.498	88.24
Age 69	3.765	54.76	4.884	77.21
Age 70	3.660	50.85	4.787	73.33
1921				
1927	−0.050	−0.43	−0.092	−0.94
1930	0.142	1.30	0.166	1.82
1933	0.142	1.22	0.212	2.23
1934	0.272	2.44	0.247	2.66
1935	0.398	3.68	0.152	1.67
1936	0.393	3.67	0.177	1.96
1937	0.286	2.70	0.050	0.56
1938	0.332	3.18	−0.034	−0.38
1939	0.428	4.15	−0.103	−1.17
1940	0.373	3.63	−0.189	−2.16
1941	0.448	4.39	−0.225	−2.58
1942	0.455	4.48	−0.276	−3.17
1943	0.497	4.90	−0.315	−3.64
1944	0.507	5.01	−0.390	−4.50
1945	0.549	5.44	−0.510	−5.90
1946	0.516	5.11	−0.691	−8.01
1947	0.560	5.57	−0.808	−9.37
1948	0.490	4.57	−1.027	−11.90
1949	0.518	5.15	−1.118	−12.92
1950	−0.065	−1.69	0.090	1.91
1951	−0.147	−3.81	−0.023	−0.49
1952	−0.154	−4.00	−0.031	−0.66
1953	−0.098	−2.57	−0.020	−0.42
1954	−0.077	−2.01	−0.046	−0.95
1955	−0.088	−2.27	−0.050	−1.01
1956	−0.024	−0.62	−0.058	−1.16
1957	−0.009	−0.22	−0.141	−2.70
1958	−0.077	−1.86	−0.055	−1.04
1959	−0.053	−1.24	−0.035	−0.63
1960	0.048	1.07	0.038	0.65
1961				

Table A.3: Extended estimation results for Table 4.1: marginal effects of the multinomial logit model for age and cohort effects (in %-points).

	Partly retired		Fully retired	
	Marginal effect	t-value	Marginal effect	t-value
Age 50				
Age 51	0.19	0.72	0.83	2.38
Age 52	0.88	3.25	1.52	4.35
Age 53	1.64	5.97	1.07	3.04
Age 54	1.97	7.12	1.53	4.36
Age 55	3.14	11.40	2.81	8.15
Age 56	4.28	15.65	4.21	12.46
Age 57	4.93	17.97	6.16	18.47
Age 58	5.98	21.81	8.19	24.84
Age 59	6.86	24.97	10.66	32.72
Age 60	6.99	25.14	16.74	52.09
Age 61	7.27	25.53	22.19	69.20
Age 62	9.07	31.07	27.20	84.18
Age 63	10.54	34.70	30.42	92.08
Age 64	10.96	34.59	32.19	95.22
Age 65	14.51	41.24	38.90	104.59
Age 66	18.17	45.18	41.39	96.64
Age 67	18.56	43.17	40.75	90.14
Age 68	18.89	39.49	42.13	84.75
Age 69	20.10	35.48	45.97	77.50
Age 70	19.36	32.36	45.19	73.74
1921				
1927	−0.15	−0.15	−0.96	−1.07
1930	0.84	0.91	1.49	1.80
1933	0.63	0.63	2.09	2.38
1934	1.93	2.02	1.96	2.30
1935	3.76	4.12	0.17	0.20
1936	3.60	3.98	0.51	0.62
1937	2.97	3.34	−0.64	−0.79
1938	3.87	4.43	−1.94	−2.40
1939	5.26	6.10	−3.26	−4.08
1940	5.03	5.86	−4.12	−5.18
1941	6.03	7.08	−4.93	−6.22
1942	6.34	7.48	−5.61	−7.11
1943	6.99	8.27	−6.31	−8.02
1944	7.44	8.83	−7.32	−9.31
1945	8.46	10.05	−9.06	−11.54
1946	8.90	10.60	−11.24	−14.34
1947	9.92	11.85	−12.94	−16.54
1948	10.12	12.08	−15.45	−19.68
1949	10.85	12.93	−16.75	−21.26
1950	−1.14	−2.49	1.46	2.37
1951	−1.54	−3.35	0.36	0.58
1952	−1.58	−3.43	0.29	0.46
1953	−1.00	−2.19	0.18	0.29
1954	−0.66	−1.42	−0.25	−0.39
1955	−0.76	−1.63	−0.25	−0.38
1956	−0.01	−0.02	−0.64	−0.98
1957	0.54	1.11	−1.78	−2.61
1958	−0.61	−1.22	−0.37	−0.53
1959	−0.43	−0.84	−0.21	−0.30
1960	0.36	0.67	0.27	0.36
1961				

Table A.4: Estimates of the multinomial logit model with fully working as the reference category, men aged 50-70 under the functional form approach.[a]

	Partly retired		Fully retired	
	Estimate	t-value	Estimate	t-value
Position in household				
Cohabiting				
Single	0.476	40.36	0.761	70.43
Other	0.491	13.04	0.641	19.24
Children				
No children				
Only children < 18	0.147	10.63	−0.278	−17.10
Both children < 18 and >=18	−0.269	−14.98	−0.637	−30.20
Only children >= 18	−0.313	−29.69	−0.511	−51.24
Education				
Primary				
Lower secondary	−0.197	−11.86	−0.646	−49.42
Higher secondary	−0.129	−8.49	−0.810	−67.69
Tertiary	0.130	8.51	−1.183	−93.29
Age				
Dummy variables	Yes	***	Yes	***
Period (year)				
Dummy variables	Yes	***	Yes	***
Cohort (year of birth)				
1/(1+exp(-(cohort-1945)))[b]	−0.163	−7.41	−0.700	−36.43
β_0	−1.940	−52.06	−1.330	−35.96
Log-likelihood	-490450.77			

[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

Table A.5: Extended estimation results for Table A.4: estimates of the multinomial logit model for age and year effects.

	Partly retired		Fully retired	
	Estimate	t-value	Estimate	t-value
Age 50				
Age 51	0.036	1.60	0.092	3.49
Age 52	0.119	5.35	0.189	7.28
Age 53	0.164	7.39	0.199	7.64
Age 54	0.205	9.23	0.285	11.12
Age 55	0.365	16.59	0.489	19.50
Age 56	0.523	23.98	0.715	29.19
Age 57	0.649	29.59	0.965	39.97
Age 58	0.824	37.31	1.240	51.66
Age 59	0.991	44.21	1.546	64.37
Age 60	1.213	51.82	2.145	88.42
Age 61	1.435	57.64	2.696	107.69
Age 62	1.796	67.18	3.272	123.81
Age 63	2.039	69.92	3.661	129.42
Age 64	2.121	67.31	3.871	129.04
Age 65	2.688	70.28	4.624	128.74
Age 66	3.114	66.75	4.991	113.96
Age 67	3.087	61.40	4.958	106.18
Age 68	3.135	55.40	5.118	98.19
Age 69	3.560	48.78	5.526	86.76
Age 70	3.215	45.19	5.450	83.99
1992	−0.441	−9.47	0.627	17.67
1993	−0.464	−9.69	0.692	19.24
1994	−0.387	−8.12	0.722	19.91
1995	−0.266	−6.05	0.770	22.14
1996	−0.344	−7.72	0.728	20.90
1997	−0.311	−7.19	0.712	20.73
1998	−0.289	−6.71	0.602	17.19
1999	−0.201	−4.90	0.809	24.25
2000	−0.020	−0.53	0.804	24.90
2001	0.062	1.78	0.689	21.71
2002	0.038	1.49	0.456	18.14
2003	0.091	4.24	0.479	22.13
2004	0.109	5.47	0.491	24.05
2005	0.094	4.85	0.498	25.25
2006	0.133	7.13	0.497	26.17
2007	0.120	6.48	0.365	19.35
2008	0.090	4.91	0.274	14.71
2009	0.018	0.99	0.137	7.29
2010	0.043	2.34	0.134	7.03
2011	0.037	2.06	0.112	6.02
2012				

Table A.6: Extended estimation results for Table 4.2: marginal effects of the multinomial logit model for age and year effects (in %-points).

	Partly retired		Fully retired	
	Marginal effect	t-value	Marginal effect	t-value
Age 50				
Age 51	−0.02	−0.06	1.03	2.97
Age 52	0.47	1.79	1.91	5.60
Age 53	0.94	3.56	1.82	5.36
Age 54	1.01	3.85	2.74	8.20
Age 55	1.88	7.28	4.65	14.24
Age 56	2.62	10.32	6.87	21.57
Age 57	2.90	11.47	9.51	30.47
Age 58	3.62	14.35	12.27	39.80
Age 59	4.11	16.27	15.47	50.55
Age 60	3.90	15.09	22.18	73.01
Age 61	3.89	14.51	28.30	92.19
Age 62	5.33	19.16	34.09	108.47
Age 63	6.29	21.30	38.02	115.99
Age 64	6.25	19.90	40.36	117.96
Age 65	9.20	25.80	47.51	123.91
Age 66	12.31	29.94	50.32	113.52
Age 67	12.16	27.59	50.01	106.79
Age 68	11.97	24.62	51.86	101.47
Age 69	12.65	22.28	56.10	93.51
Age 70	11.37	19.23	55.78	91.16
1992	−7.77	−16.08	10.08	23.79
1993	−8.33	−16.68	11.02	25.51
1994	−7.60	−15.31	11.06	25.45
1995	−6.46	−14.16	11.13	26.84
1996	−7.15	−15.37	10.95	26.12
1997	−6.70	−14.83	10.59	25.56
1998	−5.95	−13.20	9.08	21.38
1999	−5.91	−13.62	11.35	27.77
2000	−3.86	−9.78	10.46	26.69
2001	−2.41	−6.66	8.60	22.79
2002	−1.63	−6.09	5.71	18.96
2003	−1.15	−5.07	5.78	21.98
2004	−0.99	−4.69	5.84	23.51
2005	−1.20	−5.80	6.00	24.97
2006	−0.76	−3.83	5.81	25.22
2007	−0.31	−1.58	4.17	18.31
2008	−0.23	−1.23	3.12	13.97
2009	−0.42	−2.17	1.69	7.47
2010	−0.12	−0.62	1.53	6.70
2011	−0.09	−0.47	1.27	5.73
2012				

Table A.7: Estimates of the sequential multinomial logit model with partial retirement as the reference category for Stage 1 and partial retirement through self-employment as the reference category for Stage 2, men aged 50-70 under the proxy variable approach.[a]

	Fully working		Fully retired		Partly retired (self-employed)	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Position in household						
Cohabiting						
Single	−0.474	−40.13	0.289	23.37	−0.004	−0.17
Other	−0.488	−12.97	0.155	3.88	−0.092	−1.13
Children						
No children						
Only children < 18	−0.144	−10.39	−0.426	−22.47	−0.111	−3.63
Both children < 18 and >=18	0.273	15.20	−0.368	−14.43	−0.020	−0.47
Only children >= 18	0.324	30.67	−0.193	−15.60	0.150	0.64
Education						
Primary						
Lower secondary	0.258	12.13	−0.416	−19.15	−0.244	−4.97
Higher secondary	0.219	8.92	−0.632	−25.06	−0.610	−11.00
Tertiary	−0.024	−0.89	−1.253	−45.29	−0.884	−14.66
Age						
Dummy variables	Yes	***	Yes	***	Yes	***
Period (year)						
U_t	0.018	5.12	0.010	−25.73	−0.007	−0.92
Cohort (year of birth)						
Dummy variables	Yes	***	Yes	***	Yes	***
VPL	−0.610	−5.77	−2.002	2.83	0.562	3.99
β_0	2.457	22.84	2.248	31.07	1.686	11.66
Log-likelihood	-539202.93					

[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

Table A.8: Extended estimation results for Table A.7: estimates of the SMNL model with partial retirement as the reference category for Stage 1 and partial retirement through self-employment as the reference category for Stage 2, men aged 50-70 under the proxy variable approach.

	Fully working		Fully retired		Partly retired (self-employed)	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Age 50						
Age 51	−0.050	−2.23	0.032	0.95	−0.193	−3.41
Age 52	−0.148	−6.43	0.022	0.67	−0.226	−3.93
Age 53	−0.210	−9.05	−0.053	−1.58	−0.111	−1.88
Age 54	−0.261	−11.11	−0.051	−1.52	−0.338	−5.80
Age 55	−0.429	−18.27	−0.061	−1.83	−0.234	−3.96
Age 56	−0.599	−25.52	−0.062	−1.90	−0.286	−4.86
Age 57	−0.737	−31.10	−0.001	−0.04	−0.170	−2.83
Age 58	−0.920	−38.57	0.037	1.16	−0.210	−3.51
Age 59	−1.102	−45.69	0.111	3.48	−0.374	−6.30
Age 60	−1.337	−53.80	0.429	13.42	−0.435	−7.26
Age 61	−1.565	−60.21	0.704	21.83	−0.688	−11.42
Age 62	−1.934	−70.53	0.853	26.26	−0.902	−14.90
Age 63	−2.204	−75.16	0.927	28.07	−1.093	−17.79
Age 64	−2.312	−74.30	0.994	29.50	−1.159	−18.43
Age 65	−2.926	−78.61	1.116	32.12	−1.451	−22.38
Age 66	−3.398	−74.86	1.002	27.65	−1.396	−20.59
Age 67	−3.415	−69.77	0.942	25.00	−1.313	−18.56
Age 68	−3.502	−62.91	0.996	25.02	−1.555	−20.67
Age 69	−3.765	−54.76	1.119	26.14	−1.576	−19.31
Age 70	−3.660	−50.85	1.127	24.89	−1.422	−16.45
1921						
1927	0.050	0.43	−0.042	−0.60	−0.272	−1.96
1930	−0.142	−1.30	0.024	0.37	−0.129	1.00
1933	−0.142	−1.22	0.070	0.97	0.040	0.28
1934	−0.272	−2.44	−0.024	−0.36	0.180	1.34
1935	−0.398	−3.68	−0.245	−3.80	0.140	1.10
1936	−0.393	−3.67	−0.216	−3.39	0.337	2.67
1937	−0.286	−2.70	−0.235	−3.76	0.536	4.30
1938	−0.332	−3.18	−0.366	−6.00	0.424	3.50
1939	−0.428	−4.15	−0.531	−8.89	0.426	3.60
1940	−0.373	−3.63	−0.562	−9.48	0.533	4.52
1941	−0.448	−4.39	−0.673	−11.46	0.527	4.53
1942	−0.455	−4.48	−0.731	−12.51	0.712	6.14
1943	−0.497	−4.90	−0.812	−13.95	0.680	5.89
1944	−0.507	−5.01	−0.897	−15.44	0.759	6.57
1945	−0.549	−5.44	−1.060	−18.26	0.813	7.03
1946	−0.516	−5.11	−1.207	−20.85	0.760	6.61
1947	−0.560	−5.57	−1.368	−23.71	0.816	7.12
1948	−0.490	−4.87	−1.517	−26.14	0.857	7.45
1949	−0.518	−5.15	−1.636	−28.01	0.886	7.66
1950	0.065	1.69	0.155	2.71	0.317	3.44
1951	0.147	3.81	0.123	2.13	0.232	2.51
1952	0.154	4.00	0.122	2.11	0.268	2.89
1953	0.098	2.57	0.078	1.34	0.362	3.91
1954	0.077	2.01	0.031	0.53	0.321	3.43
1955	0.088	2.27	0.038	0.64	0.113	1.21
1956	0.024	0.62	−0.034	−0.56	0.104	1.11
1957	0.009	0.22	−0.132	−2.11	0.002	0.02
1958	0.077	1.86	0.021	0.33	−0.095	−0.97
1959	0.053	1.24	0.018	0.27	−0.052	−0.52
1960	−0.048	−1.07	−0.010	−0.14	0.240	2.18
1961						

Table A.9: Estimates of the multinomial logit model with being fully working as the reference category for full retirement, partly retired (1-12 hours), partly retired (13-24 hours), or partly retired (25-34 hours) (in %-points), men aged 50-70 under the proxy variable approach.[a]

	Fully retired		Partly retired (1-12 hours)		Partly retired (13-24 hours)		Partly retired (25-34 hours)	
	Estimate	t-value	Estimate	t-value	Estimate	t-value	Estimate	t-value
Position in household								
Cohabiting								
Single	0.757	69.37	0.335	13.35	0.506	27.23	0.501	33.02
Other	0.626	18.59	0.037	0.38	0.567	9.85	0.573	11.89
Children								
No children								
Only children < 18	−0.289	−17.60	−0.123	−3.24	0.187	7.90	0.169	9.76
Both children < 18 and >=18	−0.646	−30.53	−0.487	−9.38	−0.199	−6.40	−0.265	−11.73
Only children >= 18	−0.520	−51.91	−0.411	−17.39	−0.277	−16.10	−0.320	−22.82
Education								
Primary								
Lower secondary	−0.685	−38.99	−0.412	−9.98	−0.335	−10.31	−0.083	−2.72
Higher secondary	−0.865	−41.97	−0.373	−7.79	−0.332	−8.80	0.001	0.01
Tertiary	−1.269	−55.72	−0.294	−5.60	−0.251	−6.06	0.408	10.73
Age								
Dummy variables	Yes	***	Yes	***	Yes	***	Yes	***
Period (year)								
U_t	−0.009	−2.97	−0.021	−3.02	−0.014	−2.60	−0.009	−1.86
Cohort (year of birth)								
Dummy variables	Yes	***	Yes	***	Yes	***	Yes	***
VPL	−1.386	−14.41	0.601	3.75	0.283	1.95	0.606	3.90
β_0	−0.187	−2.00	−4.539	−30.83	−3.492	−24.33	−3.173	−19.85
Log-likelihood	-579038.42							

[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

Table A.10: Extended estimation results for Table A.9: estimates of the multinomial logit model for age and cohort effects.

	Fully retired		Partly retired (1-12 hours)		Partly retired (13-24 hours)		Partly retired (25-34 hours)	
	Estimate	t-value	Estimate	t-value	Estimate	t-value	Estimate	t-value
Age 50								
Age 51	0.082	3.07	0.175	2.32	0.042	1.00	0.036	1.32
Age 52	0.170	6.33	0.205	2.63	0.144	3.40	0.137	4.95
Age 53	0.157	5.81	0.323	4.18	0.282	6.72	0.156	5.52
Age 54	0.210	7.84	0.540	7.16	0.358	8.54	0.167	5.75
Age 55	0.368	13.92	0.748	10.04	0.476	11.31	0.352	12.14
Age 56	0.536	20.62	0.934	12.63	0.634	15.19	0.525	18.11
Age 57	0.735	28.59	1.005	13.47	0.739	17.56	0.694	23.79
Age 58	0.957	37.51	1.300	17.67	0.891	21.08	0.873	29.69
Age 59	1.214	47.88	1.689	23.44	1.154	27.53	0.956	31.67
Age 60	1.770	70.00	2.105	29.32	1.465	34.65	1.073	33.83
Age 61	2.274	88.60	2.532	35.26	1.712	39.73	1.171	34.47
Age 62	2.795	105.01	3.113	43.43	2.131	48.15	1.311	35.15
Age 63	3.138	112.31	3.515	48.60	2.402	52.29	1.376	32.88
Age 64	3.312	113.79	3.646	49.61	2.489	52.01	1.445	31.98
Age 65	4.046	117.17	4.466	58.72	3.036	56.84	1.662	28.57
Age 66	4.404	104.11	4.915	60.55	3.486	57.35	2.221	33.63
Age 67	4.360	96.51	5.052	60.69	3.372	51.07	2.106	28.17
Age 68	4.501	88.27	5.119	57.96	3.456	47.11	2.252	26.92
Age 69	4.887	77.23	5.339	54.16	3.761	43.71	2.570	26.15
Age 70	4.789	73.35	5.332	52.79	3.525	37.86	2.383	22.04
1921								
1927	−0.093	−0.95	−0.009	−0.06	−0.034	−0.22	−0.228	−1.24
1930	0.162	1.78	0.173	1.36	0.126	0.89	0.005	0.03
1933	0.212	2.23	0.281	2.04	0.115	0.74	−0.151	−0.83
1934	0.252	2.71	0.436	3.33	0.298	2.03	−0.118	−0.68
1935	0.156	1.71	0.497	3.93	0.467	3.32	0.085	0.51
1936	0.184	2.03	0.477	3.81	0.575	4.16	−0.063	−0.38
1937	0.049	0.55	0.264	2.12	0.434	3.18	0.042	0.26
1938	−0.026	−0.29	0.502	4.16	0.336	2.48	0.021	0.13
1939	−0.099	−1.13	0.490	4.10	0.540	4.06	0.136	0.87
1940	−0.185	−2.11	0.442	3.73	0.484	3.66	0.065	0.42
1941	−0.222	−2.55	0.495	4.20	0.555	4.22	0.175	1.13
1942	−0.270	−3.11	0.499	4.25	0.629	4.81	0.102	0.66
1943	−0.312	−3.60	0.566	4.83	0.565	4.33	0.255	1.67
1944	−0.388	−4.48	0.564	4.83	0.552	4.24	0.300	1.98
1945	−0.508	−5.88	0.604	5.17	0.591	4.55	0.346	2.28
1946	−0.692	−8.01	0.498	4.27	0.571	4.40	0.347	2.29
1947	−0.807	−9.36	0.552	4.75	0.607	4.69	0.392	2.60
1948	−1.029	−11.92	0.400	3.42	0.472	3.64	0.404	2.67
1949	−1.121	−12.97	0.325	2.75	0.487	3.75	0.465	3.08
1950	0.079	1.68	−0.345	−2.91	0.110	1.51	−0.037	−0.79
1951	−0.031	−0.65	−0.521	−4.32	0.051	0.70	−0.128	−2.75
1952	−0.036	−0.75	−0.387	−3.21	−0.007	−0.10	−0.146	−3.16
1953	−0.024	−0.49	−0.305	−2.53	0.042	0.57	−0.100	−2.18
1954	−0.049	−1.01	−0.354	−2.88	0.054	0.74	−0.069	−1.50
1955	−0.052	−1.06	−0.167	−1.36	0.070	0.95	−0.121	−2.58
1956	−0.061	−1.21	−0.329	−2.61	0.086	1.14	−0.005	−0.10
1957	−0.144	−2.74	−0.076	−0.60	0.041	0.53	0.002	0.04
1958	−0.057	−1.07	−0.158	−1.20	0.042	0.53	−0.094	−1.89
1959	−0.036	−0.64	−0.252	−1.80	−0.075	−0.89	−0.012	−0.24
1960	0.037	0.65	−0.008	−0.06	0.018	0.20	0.068	1.29
1961								

Table A.11: Extended estimation results for Table 4.3: marginal effects of the multinomial logit model for age and cohort effects (in %-points).

	Fully retired		Partly retired (1-12 hours)		Partly retired (13-24 hours)		Partly retired (25-34 hours)	
	Marginal effect	t-value	Marginal effect	t-value	Marginal effect	t-value	Marginal effect	t-value
Age 50								
Age 51	0.68	1.88	0.33	1.63	0.00	0.02	0.05	0.31
Age 52	1.45	4.00	0.25	1.20	0.24	1.38	0.47	2.91
Age 53	0.86	2.36	0.56	2.70	0.80	4.61	0.53	3.27
Age 54	1.09	3.02	1.04	5.15	0.99	5.71	0.46	2.77
Age 55	2.37	6.65	1.31	6.57	1.13	6.49	1.22	7.36
Age 56	3.76	10.76	1.50	7.61	1.42	8.26	1.89	11.42
Age 57	5.80	16.80	1.35	6.82	1.47	8.47	2.53	15.22
Age 58	7.73	22.65	1.75	8.96	1.63	9.41	3.14	18.83
Age 59	9.93	29.47	2.35	12.26	2.20	12.86	3.10	18.25
Age 60	15.81	47.63	2.56	13.48	2.47	14.44	2.82	15.95
Age 61	21.13	63.85	2.89	15.28	2.60	15.00	2.52	13.46
Age 62	26.12	78.35	3.58	19.03	3.31	18.95	2.36	11.65
Age 63	29.42	86.40	4.08	21.67	3.77	21.14	2.10	9.30
Age 64	31.23	89.70	4.15	21.79	3.81	20.66	2.20	9.05
Age 65	38.29	100.04	5.12	26.64	4.66	23.68	2.10	6.79
Age 66	40.72	93.01	5.62	28.46	5.68	26.62	4.56	13.36
Age 67	40.30	87.11	6.08	30.34	5.30	22.90	3.98	10.21
Age 68	41.67	82.06	6.01	29.07	5.36	21.23	4.58	10.56
Age 69	45.38	75.36	5.93	27.24	5.86	20.69	5.71	11.43
Age 70	44.77	71.81	6.12	27.52	5.11	16.32	4.86	8.72
1921								
1927	−0.81	−0.90	0.15	0.58	0.08	0.16	−1.15	−1.19
1930	1.62	1.94	0.20	0.79	0.22	0.46	−0.27	−0.31
1933	2.35	2.65	0.43	1.54	0.13	0.23	−1.24	−1.29
1934	2.29	2.65	0.75	2.87	0.78	1.53	−1.19	−1.30
1935	0.40	0.47	1.00	3.98	1.55	3.22	0.04	0.05
1936	0.82	0.99	0.91	3.67	1.99	4.23	−0.87	−1.00
1937	−0.52	−0.63	0.55	2.24	1.64	3.52	0.00	0.00
1938	−1.65	−2.03	1.31	5.56	1.33	2.89	−0.02	−0.03
1939	−3.07	−3.81	1.34	5.76	2.25	4.99	0.68	0.84
1940	−3.91	−4.88	1.36	5.89	2.19	4.89	0.42	0.52
1941	−4.75	−5.95	1.53	6.68	2.50	5.62	1.07	1.34
1942	−5.39	−6.78	1.61	7.06	2.90	6.57	0.70	0.88
1943	−6.14	−7.75	1.83	8.07	2.65	6.01	1.64	2.07
1944	−7.16	−9.04	1.93	8.52	2.71	6.15	2.01	2.55
1945	−8.90	−11.25	2.20	9.72	3.04	6.92	2.42	3.08
1946	−11.07	−14.02	2.20	9.68	3.27	7.46	2.71	3.46
1947	−12.76	−16.19	2.50	11.05	3.59	8.21	3.11	3.98
1948	−15.20	−19.22	2.44	10.67	3.42	7.80	3.56	4.55
1949	−16.39	−20.64	2.37	10.17	3.62	8.24	4.05	5.18
1950	1.41	2.23	−1.04	−3.26	0.38	1.26	−0.30	−1.13
1951	0.47	0.74	−1.32	−4.07	0.37	1.23	−0.62	−2.32
1952	0.33	0.51	−0.94	−2.92	0.13	0.42	−0.72	−2.37
1953	0.22	0.34	−0.76	−2.35	0.28	0.93	−0.50	−1.89
1954	−0.10	−0.16	−0.86	−2.60	0.37	1.22	−0.29	−1.08
1955	−0.37	−0.56	−0.35	−1.08	0.43	1.40	−0.61	−2.25
1956	−0.43	−0.64	−0.79	−2.33	0.50	1.59	0.09	0.31
1957	−1.81	−2.59	0.01	0.02	0.41	1.27	0.21	0.77
1958	−0.43	−0.61	−0.32	−0.92	0.31	0.95	−0.44	−1.54
1959	0.05	0.07	−0.60	−1.59	−0.21	−0.59	0.04	0.15
1960	0.38	0.48	−0.09	−0.23	−0.01	−0.02	0.33	1.07
1961								

Table A.12: Extended estimation results for Table 4.6: estimates of the multinomial probit model for age and cohort effects under different starting values of ρ and σ_{33} for partial retirement.[a]

	Model (1)		Model (2)		Model (3)	
	Estimate	std. err.	Estimate	std. err.	Estimate	std. err.
Age 50						
Age 51	0.026	0.002521	0.026	0.001793	0.026	0.015519
Age 52	0.074	0.001432	0.074	0.001174	0.074	0.010250
Age 53	0.105	0.003611	0.105	0.001794	0.105	0.014515
Age 54	0.133	0.003236	0.133	0.001447	0.133	0.018928
Age 55	0.219	0.001732	0.219	0.001052	0.219	0.012292
Age 56	0.309	0.002597	0.309	0.001046	0.309	0.013019
Age 57	0.382	0.001964	0.382	0.000767	0.382	0.016219
Age 58	0.479	0.001497	0.479	0.001141	0.479	0.017330
Age 59	0.573	0.002405	0.573	0.001844	0.574	0.009353
Age 60	0.684	0.003197	0.684	0.001410	0.685	0.012863
Age 61	0.781	0.002424	0.780	0.000744	0.781	0.012154
Age 62	0.942	0.001653	0.942	0.002711	0.943	0.012359
Age 63	1.049	0.001627	1.048	0.004708	1.049	0.013547
Age 64	1.079	0.004101	1.079	0.002739	1.080	0.012715
Age 65	1.299	0.004336	1.299	0.001651	1.298	0.014955
Age 66	1.486	0.002770	1.486	0.001285	1.484	0.016635
Age 67	1.495	0.004951	1.495	0.002315	1.494	0.016314
Age 68	1.511	0.002297	1.512	0.001492	1.510	0.021330
Age 69	1.578	0.002942	1.579	0.002814	1.576	0.015572
Age 70	1.539	0.003423	1.540	0.002903	1.537	0.019315
1921						
1927	−0.004	0.001415	−0.004	0.001760	−0.004	0.022139
1930	0.060	0.001451	0.060	0.000821	0.059	0.016341
1933	0.062	0.004028	0.062	0.002501	0.061	0.026922
1934	0.137	0.001908	0.137	0.002332	0.136	0.023244
1935	0.201	0.002556	0.202	0.001968	0.201	0.024139
1936	0.205	0.002233	0.205	0.000983	0.204	0.017605
1937	0.156	0.004093	0.157	0.001215	0.156	0.019875
1938	0.190	0.002303	0.190	0.000888	0.189	0.023852
1939	0.249	0.002154	0.249	0.006089	0.247	0.019363
1940	0.227	0.002304	0.227	0.003615	0.226	0.015864
1941	0.271	0.001205	0.271	0.001442	0.269	0.016274
1942	0.277	0.001219	0.278	0.002558	0.276	0.013061
1943	0.303	0.001886	0.304	0.000792	0.302	0.012700
1944	0.313	0.002208	0.313	0.005423	0.311	0.014268
1945	0.342	0.002850	0.342	0.001888	0.340	0.012808
1946	0.334	0.003095	0.335	0.000552	0.333	0.015013
1947	0.362	0.001529	0.363	0.002139	0.360	0.017779
1948	0.330	0.003585	0.331	0.001782	0.328	0.011364
1949	0.345	0.004385	0.345	0.001505	0.342	0.014362
1950	−0.017	0.003798	−0.017	0.001945	−0.017	0.009196
1951	−0.064	0.001235	−0.064	0.000887	−0.064	0.012284
1952	−0.069	0.002344	−0.069	0.001141	−0.069	0.009474
1953	−0.040	0.003544	−0.040	0.001481	−0.040	0.010879
1954	−0.031	0.001720	−0.031	0.001417	−0.031	0.014102
1955	−0.037	0.002538	−0.037	0.000743	−0.037	0.010698
1956	−0.005	0.002187	−0.005	0.000996	−0.005	0.014485
1957	0.003	0.002225	0.003	0.001359	0.003	0.013073
1958	−0.033	0.003967	−0.033	0.000976	−0.033	0.013232
1959	−0.024	0.002123	−0.024	0.002436	−0.024	0.014669
1960	0.028	0.003092	0.028	0.004831	0.028	0.016361
1961						

[a] The starting values for Model (1) are: $\rho = 0$ and $\sigma_{33} = 1.00$, Model (2): $\rho = 0.60$ and $\sigma_{33} = 1.50$ and Model (3): $\rho = -0.80$, and $\sigma_{33} = 0.25$.

Table A.13: Extended estimation results for Table 4.6: estimates of the multinomial probit model for age and cohort effects under different starting values of ρ and σ_{33} for full retirement.[a]

	Model (1)		Model (2)		Model (3)	
	Estimate	std. err.	Estimate	std. err.	Estimate	std. err.
Age 50						
Age 51	0.037	0.002533	0.038	0.003511	0.037	0.010986
Age 52	0.079	0.002278	0.079	0.000711	0.078	0.009992
Age 53	0.074	0.001357	0.075	0.000650	0.074	0.013066
Age 54	0.099	0.001242	0.100	0.000946	0.098	0.009723
Age 55	0.174	0.002116	0.175	0.000499	0.172	0.008311
Age 56	0.258	0.001443	0.258	0.000562	0.255	0.008923
Age 57	0.356	0.001553	0.357	0.001345	0.353	0.011433
Age 58	0.470	0.003756	0.471	0.001181	0.465	0.009978
Age 59	0.603	0.001459	0.604	0.002216	0.597	0.009252
Age 60	0.896	0.001318	0.898	0.001680	0.887	0.009034
Age 61	1.164	0.002133	1.167	0.001447	1.152	0.008793
Age 62	1.432	0.001980	1.436	0.001824	1.417	0.010958
Age 63	1.597	0.002018	1.602	0.001786	1.581	0.010261
Age 64	1.678	0.002816	1.683	0.000955	1.661	0.009544
Age 65	1.993	0.004751	1.999	0.002426	1.973	0.011324
Age 66	2.125	0.003261	2.131	0.003797	2.104	0.014274
Age 67	2.100	0.004688	2.107	0.002330	2.080	0.013857
Age 68	2.150	0.008730	2.156	0.005480	2.129	0.017022
Age 69	2.290	0.005377	2.297	0.001945	2.268	0.017120
Age 70	2.254	0.010912	2.261	0.005795	2.231	0.020190
1921						
1927	−0.036	0.004734	−0.036	0.001754	−0.035	0.016515
1930	0.070	0.001149	0.070	0.000552	0.070	0.012412
1933	0.093	0.003940	0.093	0.001378	0.092	0.015753
1934	0.118	0.002168	0.118	0.002768	0.117	0.013851
1935	0.056	0.001786	0.056	0.004117	0.056	0.017737
1936	0.074	0.002594	0.074	0.002846	0.074	0.012672
1937	0.007	0.002250	0.006	0.003031	0.007	0.013255
1938	−0.037	0.001938	−0.037	0.000865	−0.036	0.013648
1939	−0.073	0.005665	−0.074	0.000805	−0.071	0.013222
1940	−0.114	0.002841	−0.115	0.003607	−0.112	0.011494
1941	−0.135	0.002234	−0.136	0.002024	−0.133	0.011657
1942	−0.163	0.002867	−0.164	0.000803	−0.160	0.013554
1943	−0.185	0.002837	−0.186	0.000751	−0.181	0.014855
1944	−0.225	0.003705	−0.226	0.001052	−0.221	0.016561
1945	−0.292	0.001440	−0.293	0.001130	−0.287	0.011952
1946	−0.386	0.001899	−0.388	0.001600	−0.380	0.012302
1947	−0.451	0.002405	−0.453	0.001200	−0.444	0.011754
1948	−0.564	0.002366	−0.567	0.000918	−0.556	0.011833
1949	−0.610	0.002283	−0.613	0.001083	−0.601	0.011415
1950	0.025	0.001349	0.026	0.003885	0.025	0.010803
1951	−0.022	0.002435	−0.022	0.001153	−0.022	0.012821
1952	−0.026	0.003571	−0.026	0.004437	−0.026	0.012510
1953	−0.018	0.002754	−0.018	0.003296	−0.018	0.011665
1954	−0.028	0.001908	−0.028	0.004550	−0.027	0.012826
1955	−0.027	0.002758	−0.027	0.003419	−0.027	0.013448
1956	−0.029	0.002457	−0.030	0.003350	−0.029	0.012727
1957	−0.063	0.002462	−0.064	0.003159	−0.063	0.014617
1958	−0.029	0.003772	−0.029	0.001404	−0.029	0.014998
1959	−0.019	0.002973	−0.019	0.001763	−0.019	0.018424
1960	0.021	0.003227	0.021	0.002186	0.021	0.018112
1961						

[a] The starting values for Model (1) are: $\rho = 0$ and $\sigma_{33} = 1.00$, Model (2): $\rho = 0.60$ and $\sigma_{33} = 1.50$ and Model (3): $\rho = -0.80$, and $\sigma_{33} = 0.25$.

Table A.14: Estimates of the multinomial logit model with fully working as the reference category, men aged 50-70 under the proxy variable approach with only 50% of the observations.[a,b]

	Partly retired		Fully retired	
	Estimate	t-value	Estimate	t-value
Position in household				
Cohabiting				
Single	0.478	28.65	0.763	49.77
Other	0.418	7.71	0.568	11.93
Children				
No children				
Only children < 18	0.150	7.68	−0.279	−12.07
Both children < 18 and >=18	−0.304	−11.92	−0.638	−21.49
Only children >= 18	−0.339	−22.67	−0.509	−36.25
Education				
Primary				
Lower secondary	−0.267	−8.90	−0.679	−27.53
Higher secondary	−0.219	−6.31	−0.849	−29.30
Tertiary	0.005	0.14	−1.239	−38.78
Age				
Dummy variables	Yes	***	Yes	***
Period (year)				
U_t	−0.019	−3.79	−0.008	−1.85
Cohort (year of birth)				
Dummy variables	Yes	***	Yes	***
VPL	0.710	4.91	−1.341	−10.22
β_0	−2.538	−17.22	−0.276	−2.19
Log-likelihood	-245677.39			

[a] Estimation results marked with *** (only used if standard errors are not given) are jointly significant at a 1% significance level (F-test).

[b] A value between zero and one is randomly assigned to each observation. Next, the sample is restricted to observations with values less than 0.50.

Table A.15: Extended estimation results for Table A.14: estimates of the multinomial logit model for age and cohort effects with only 50% of the observations.

	Partly retired		Fully retired	
	Estimate	t-value	Estimate	t-value
Age 50				
Age 51	0.051	1.61	0.054	1.44
Age 52	0.148	4.57	0.161	4.28
Age 53	0.216	6.62	0.166	4.39
Age 54	0.272	8.22	0.215	5.74
Age 55	0.448	13.52	0.362	9.75
Age 56	0.586	17.67	0.491	13.43
Age 57	0.760	22.77	0.717	19.89
Age 58	0.931	27.63	0.949	26.53
Age 59	1.127	33.15	1.195	33.58
Age 60	1.334	37.87	1.747	49.17
Age 61	1.569	42.75	2.262	62.81
Age 62	1.930	49.90	2.748	73.62
Age 63	2.217	53.46	3.112	79.17
Age 64	2.301	52.60	3.250	79.75
Age 65	2.911	55.21	4.010	82.44
Age 66	3.412	53.05	4.376	73.13
Age 67	3.325	49.04	4.252	68.47
Age 68	3.582	44.53	4.542	61.49
Age 69	3.661	38.28	4.798	55.07
Age 70	3.641	36.03	4.744	51.94
1921				
1927	0.118	0.73	0.071	0.53
1930	0.294	1.97	0.312	2.54
1933	0.270	1.69	0.325	2.52
1934	0.239	1.56	0.239	1.91
1935	0.464	3.13	0.243	1.97
1936	0.532	3.63	0.293	2.40
1937	0.406	2.80	0.150	1.25
1938	0.388	2.72	0.064	0.54
1939	0.536	3.81	−0.012	−0.11
1940	0.472	3.37	−0.109	−0.92
1941	0.489	3.51	−0.163	−1.39
1942	0.522	3.76	−0.185	−1.59
1943	0.595	4.30	−0.212	−1.82
1944	0.580	4.20	−0.300	−2.58
1945	0.656	4.76	−0.392	−3.38
1946	0.628	4.57	−0.606	−5.23
1947	0.658	4.80	−0.732	−6.32
1948	0.573	4.18	−0.929	−8.01
1949	0.610	4.44	−1.013	−8.73
1950	−0.055	−1.01	0.116	1.70
1951	−0.162	−2.95	0.030	0.43
1952	−0.165	−3.02	0.014	0.21
1953	−0.083	−1.54	0.033	0.48
1954	−0.096	−1.76	−0.023	−0.32
1955	−0.119	−2.15	−0.003	−0.04
1956	0.013	0.24	−0.037	−0.51
1957	−0.004	−0.07	−0.112	−1.49
1958	−0.044	−0.76	−0.026	−0.34
1959	−0.012	−0.20	0.007	0.09
1960	0.033	0.51	0.053	0.63
1961				

Appendix B

Test of the IIA assumption

The Independence of Irrelevant Alternatives (IIA) is a key property of the multinomial logit model. It indicates that the ratio between alternative A and alternative B should not change when a third alternative, C, is introduced. To test this property the Hausman and McFadden Test, or the Small and Hsiao Test are commonly conducted. Notice that under IIA, estimates from the full model, $\hat{\beta}_j^f$, are consistent and efficient, while estimates from the restricted model, $\hat{\beta}_j^r$ are consistent but inefficient. In this chapter, the full model is (unless stated otherwise) the Multinomial logit model (see Section 4.1) with three alternatives: fully working, gradually retired, or fully retired. The restricted model is the (Multinomial) logit model restricted to two categories rather than the three categories of the full model.

Hausman and McFadden Test

Hausman and McFadden (1984) propose a Hausman test that compares the estimates of $\hat{\beta}_j^f$ to $\hat{\beta}_j^r$. The test statistic is defined as

$$HM = (\hat{\beta}_j^r - \hat{\beta}_j^f)^T[\widehat{V}(\hat{\beta}_j^r) - \widehat{V}(\hat{\beta}_j^f)]^{-1}(\hat{\beta}_j^r - \hat{\beta}_j^f)$$

where $\widehat{V}(\hat{\beta}_j^r)$ and $\widehat{V}(\hat{\beta}_j^f)$ are the estimated covariance matrices. If IIA holds, HM is asymptotically chi-squared distributed with k degrees of freedom where k is equal to the number of coefficients estimated in the restricted model. If

$$HM > \chi^2_{0.95}(k),$$

thus rejecting the null hypothesis, the IIA assumption is unlikely to hold. Hausman and McFadden (1984) indicate that HM can be negative if $\widehat{V}(\hat{\beta}_j^r) - \widehat{V}(\hat{\beta}_j^f)$ is not positive semi-definite which they conclude "does not impede calculation of the statistic or carrying out the asymptotic test". A negative test statistic can thus be interpreted as evidence that the IIA assumption is likely to hold.

Small and Hsiao Test

Another way to test the IIA assumption was proposed by Small and Hsiao (1985). To obtain the test statistic, SH, first the sample is randomly divided into sub samples A and B of roughly equal size. The model then is estimated on both sub samples yielding estimates $\hat{\beta}^f_{A,j}$ and $\hat{\beta}^f_{B,j}$. Next a weighted average of these coefficients[1] is taken as follows

$$\hat{\beta}^f_{AB,j} = (\frac{1}{\sqrt{2}})\hat{\beta}^f_{A,j} + (1 - \frac{1}{\sqrt{2}})\hat{\beta}^f_{B,j}.$$

The restricted sub sample is created from sub sample B by eliminating all cases with a given value of the dependent variable, e.g. all individuals that are partly retired, which leads to the estimates $\hat{\beta}^r_{B,j}$ with (log-) likelihood function L. The test statistic is defined as

$$SH = -2[L(\hat{\beta}^f_{AB,j}) - L(\hat{\beta}^r_{B,j})]$$

which is asymptotically chi-squared distributed with k degrees of freedom where k is equal to the number of parameters in the restricted sub sample. Again, if the test statistic is larger than the critical value, thus rejecting the null hypotheses, the IIA assumption is unlikely to hold.

Compared to other studies that model the decision to gradually retire (see, e.g., Cahill et al. (2015)), this study contains a lot of observations. A direct consequence of having a lot of observations is that the confidence intervals of the coefficients become very narrow. Thus rejecting the null hypothesis of the two coefficients being equal fairly soon, even if the difference of the estimates is in absolute values small. Although the estimates could statistically differ from each other, one may argue whether these estimates are different from an economic point of view. Thus, while the null hypothesis of the IIA assumption to hold is statistically unlikely to hold, from an economic point of view the IIA assumption could approximately hold.

To examine this conjecture, a value between zero and one is randomly assigned to each observation. Next, the sample is restricted to observations with value, e.g., less than 0.50. Table A.14 and Table A.15 presents the estimation results for a MNL model with three categories[2] with only 50% of the observations. Although these estimates differ statistically from the estimates of the same MNL model with the complete dataset, the conclusion derived via, and interpretation of, the estimates of these two sets of results are similar. Thus, from an economic point of view, the estimation results

[1] The choice of weights $1/\sqrt{2}$ is merely technical. To proof asymptomatic convergence of the test statistic to a chi-squared distribution with k degrees of freedom, the authors use that $(1-b)^2 + (b^2 - 2b) = 0$ where $b = 1 - 1/\sqrt{2}$.

[2] The remainder of this paragraph also holds for the MNL model with five categories.

with only 50% of the observations[3] are similar to those with the complete dataset.

Table B.1 presents (see Table B.2 for MNL model with five categories of Section 4.3) the values of the HM and SH test statistics for (randomly restricted) samples. The SH test does not reject the null hypothesis (of the IIA to hold) in favour of the alternative when randomly 50%, or more, of the total number of observations is not taken into account. The HM test rejects the null hypothesis in favour of the alternative hypothesis for most of the investigated restricted samples. As expected beforehand, the null hypotheses of both tests are rejected in favour of the alternative hypothesis under the complete dataset.

In addition, Monte Carlo simulations suggest that even in well specified models these IIA tests often reject the assumption when the alternatives seem distinct (see Cheng and Long (2007)). McFadden (1973) wrote that the multinomial logit model should only be used where the outcome categories "can plausibly be assumed to be distinct and weighed independently in the eyes of each decision maker." As the choices to either gradual retire, fully work, or to fully retire are most likely distinct and can thus be weighed independently, the IIA assumption is assumed to approximately holds in the analysis of the current study. If the IIA fails to hold, a more flexible model such as a nestled logit, random parameter logit or a multinomial probit could be more appropriate (see also Section 3.2.4 of this study).

[3]This preceding line of reasoning fails to hold starting from less than 25% of the observations of the complete dataset.

Table B.1: Values of the HM and SH test statistics under both the proxy variable approach and the functional form approach.[a,b,c,d]

'Proxy variable approach'

# Obs.	HM			SH		
	FW	PR	FR	FW	PR	FR
702879	−1345.49	401.96	753.75	91.18	85.77	114.73
		***	***	***	**	***
351246	129.91	198.81	112.41	77.01	65.61	81.87
	***	***	***	*		**
175731	599.08	119.52	124.52	78.93	57.16	66.34
	***	***	***	*		

'Functional form approach'

# Obs.	HM			SH		
	FW	PR	FR	FW	PR	FR
702879	521.10	362.76	26.33	73.22	61.54	115.42
	***	***		**		***
351246	386.58	235.81	172.32	39.71	69.13	71.23
	***	***	***		*	*
175731	233.46	131.00	−7.66	66.89	61.98	50.98
	***	***		**		

[a] Estimation result marked with *, **, or *** is significant at a 10%, 5%, or 1% significance level.

[b] Under the proxy variable approach the number of degrees of freedom, k, is 61, and under the functional form approach the number of degrees of freedom is 50.

[c] FW, PR, or FR indicates to which category the restricted model is restricted. For example, PR denotes that the restricted model consists of the categories FW and FR.

[d] FW is an abbreviation for fully working, PR for partly retired, and FR for fully retired.

Table B.2: Values of the HM and SH test statistics under the proxy variable approach.[a,b,c,d]

			HM		
# Obs.	FW	PR (25-34 hrs.)	PR (13-24 hrs.)	PR (1-12 hrs.)	FR
702879	205.11	391.15 ***	171.03	39.16	−692.61
351246	−178.79	699.57 ***	152.36	99.75	−72.88
175731	647.29 ***	162.83	43.39	−92.08	−114.62

			SH		
# Obs.	FW	PR (25-34 hrs.)	PR (13-24 hrs.)	PR (1-12 hrs.)	FR
702879	217.05 **	173.09	164.45	172.13	261.69 ***
351246	206.98	189.48	170.77	196.74	227.03 **
175731	193.68	164.56	182.92	167.97	187.27

[a] Estimation result marked with *, **, or *** is significant at a 10%, 5%, or 1% significance level.

[b] The number of degrees of freedom, k, is 183.

[c] FW, PR (25-34 hrs.), PR (13-24 hrs.), PR (1-12 hrs.) and FR indicates to which category the restricted model is restricted. For example, FR denotes that the restricted model consists of the categories PR (25-34 hrs.), PR (13-24 hrs.), PR (1-12 hrs.) and FW.

[d] FW is an abbreviation for fully working, PR for partly retired, and FR for fully retired.

Appendix C

Test of the Proportionality assumption

The Proportionality assumption (sometimes referred to as either 'parallel regression assumption' and, for the ordered logit model, the 'proportional odds assumption') is key to a valid application of the ordered logit/probit model. The proportional odds assumption indicates that logarithm of the odds form an arithmetic sequence. In order to see this, observe that for the ordered choice model of Section 3.2.2, under the assumption that ϵ is standard logistic distributed, the cumulative distribution probabilities $\gamma_{i,j} = P(y_i \leq j|\mu_i(x_{it}, a_{it}, t, c_i))$ can be written as (see, e.g., Brant (1990))

$$\log(\gamma_{i,j}/(1-\gamma_{i,j})) = \alpha_j - \theta'\mu_i \tag{C.1}$$

where the parameters are in correspondence of Section 3.2.2. Notice that for each category, j, the parameter estimates θ are the same.

The common values of θ provide a possibility to test the Proportionality assumption. Augment model (3.3) by incorporating separate θ_j's, consequently equation (C.1) changes to

$$\log(\gamma_{i,j}/(1-\gamma_{i,j})) = \alpha_j - \theta_j'\mu_i \tag{C.2}$$

and apply methods such as likelihood ratio (LR) test, or a 'Brant' test to asses the hypothesis

$$H_0: \quad \theta_j = \theta, \qquad j = 1, \ldots, J-1. \tag{C.3}$$

(approximate) LR test

In general, the LR test compares the (log) likelihood of the null model with the (log) likelihood of the alternative model (see, e.g. Cameron and Trivedi (2005) p. 234). The test statistic is defined as

$$LR = -2[L(\hat{\beta}_j^1) - L(\hat{\beta}_j^2)]$$

where the estimates from the null model are defined by $\hat{\beta}_j^1$, the estimates for the alternative model by $\hat{\beta}_j^2$, and L represents the (log) likelihood function. LR is approximately chi-squared distributed with k degrees of freedom where k is the difference between the degrees of freedom of the alternative model and the degrees of freedom of the null model.

In this chapter, the log likelihood of the ordered logit model (null model) is compared with the log likelihood obtained from pooling $J - 1$ binary logit models (alternative model). The technical details of computing the approximate LR test in Stata can be found in Wolfe and Gould (1998). If

$$LR > \chi^2_{0.95}(k),$$

the null hypothesis (see (C.3)) that the Proportionality assumption holds is rejected in favour of the alternative hypothesis.

Brant test

The LR test does not distinguish between (combinations of) estimated coefficients. To test whether the estimated coefficients for some variables are identical under an ordered logit model a Brant test could be conducted. Essentially, the Brant test is a Wald test, and the details of computing this test statistic, 'Brant', can be found in Brant (1990).

For this study it is only of interest to asses the hypothesis corresponding the estimates of θ_j of the augmented model (see (C.2)) and the null model (see (C.1)) simultaneously. Consequently, if

$$\text{'Brant'} > \chi^2_{0.95}(k),$$

the null hypothesis (see (C.3)) that the Proportionality assumption holds is rejected in favour of the alternative hypothesis.

In Table C.1 the values of the LR and 'Brant' test statistics are presented for the model of Section 4.1. Table C.2 consists of the LR and 'Brant' test statistics of the model in Section 4.3. In accordance with Appendix B the values of the test statistics are presented for different (randomly restricted) samples. The null hypothesis of both tests for both models is rejected in favour of the alternative hypothesis for all of the investigated sample sizes. This suggests that a more general model, such as the multinomial logit model, is more appropriate for the data than an ordered logit model.

Table C.1: Values of the LR and 'Brant' test statistics under both the proxy variable approach and the functional form approach for $J = 3$.[a,b,c]

'Proxy variable approach'		
# Obs.	LR	'Brant'
702879	15 967.88 ***	15 357.64 ***
351246	7950.41 ***	7721.64 ***
175731	3984.22 ***	3824.82 ***

'Functional form approach'		
# Obs.	LR	'Brant'
702879	16 350.66 ***	16 566.39 ***
351246	8072.13 ***	8224.80 ***
175731	3982.48 ***	4012.93 ***

[a] Estimation result marked with *** is significant at a 1% significance level.
[b] Under the proxy variable approach the number of degrees of freedom, k, is 60, and under the functional form approach the number of degrees of freedom is 49.

Table C.2: Values of the LR and 'Brant' test statistics under the proxy variable approach where $J = 5$. [a,b,c]

'Proxy variable approach'		
# Obs.	LR	'Brant'
702879	25 778.42 ***	23 166.93 ***
351246	12 793.82 ***	11 591.87 ***
175731	6415.65 ***	5834.22 ***

[a] Estimation result marked with *** is significant at a 1% significance level.
[b] The number of degrees of freedom, k, is 180.

Appendix D

Matlab code

This appendix consists of self-written code in Matlab to perform maximum likelihood estimation.

Multinomial probit estimates

```
function [nll] = nll_MNP(pars, I, X, W, M)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This function returns the negative log-likelihood under a multinomial   %
% probit model where correlation in the error terms between different     %
% categories is possible. This particular function is written assuming    %
% that the first, of in total three, categories is the base category. In  %
% addition, sigma_22 is assumed to equal 1. For a description of the model%
% to be estimated see Section 4.4 .                                       %
%                                                                         %
% Author: J.L.M. Bonekamp                                                 %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% calls:                                                                  %
% bvnormcdf.m by Alan Genz (see http://www.maths.lth.se/matstat/wafo/docum%
%              entation/wafodoc/wafo/trgauss/bvnormcdf.html)              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% output:                                                                 %
% nll     negative sum of the log-likelihood                              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% pars    coefficient for estimation                                      %
% I       (row-sorted) element of matrix equals 1 when individual (row) is%
%         observed in that category (column) and zero otherwise           %
% X       (row-sorted) matrix of covariates                               %
% W       matrix of weights to speed up optimization                      %
% M       (only when I and X are sorted) element of matrix M indicates    %
%         which row                                                       %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
[m, n] = size(X);

rho = pars(1);  % corr. eps_2 and eps_3
```

```
sigma_33 = pars(2);
sigma_23 = rho*sqrt(sigma_33);

% [corr(eps_2, eps_3), corr(eps_2, eps_2-eps_3), corr(eps_3, eps_3-eps_2)]
RHO   = [sigma_23/sqrt(sigma_33), ...
        (1-sigma_23)/sqrt(sigma_33+1-2*sigma_23), ...
        (sigma_33-sigma_23)/(sqrt(sigma_33+1-2*sigma_23)*sqrt(sigma_33))];

pars(1:2,:) = [];
pars = reshape(pars,n,2);

V   = [X*pars(:,1) X*pars(:,2)];

% upper/under integral limit
B1 = full(sum(I.*[-V(:,1), ...
                   V(:,1), ...
                   V(:,2)/sqrt(sigma_33)],2));

B2 = full(sum(I.*[-V(:,2)/sqrt(sigma_33), ...
                  -(V(:,2)-V(:,1))/(sqrt(1+sigma_33-2*sigma_23)), ...
                  -(V(:,1)-V(:,2))/(sqrt(1+sigma_33-2*sigma_23))],2));

P = zeros(m,1); % probabilities
P(1:(M(2)-1))    =  bvnormcdf(B1(1:(M(2)-1)),   B2(1:(M(2)-1))   ,RHO(1));
P(M(2):(M(3)-1)) =  bvnormcdf(B1(M(2):(M(3)-1)),B2(M(2):(M(3)-1)),RHO(2));
P(M(3):end)      =  bvnormcdf(B1(M(3):end),     B2(M(3):end)     ,RHO(3));

P = max(P, eps);

nll = -sum(log(P).*W); % negative log-likelihood
```

Misclassification probabilities

```
function [nll] = nll_b_mlogit_1_alpha(pars, I, X)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This function returns the negative log-likelihood under a multinomial   %
% logit model with unknown misclassification probabilities. This          %
% particular function is written assuming that the first, of in total     %
% three, categories is the base category. In addition, only the           %
% probability of misclassification of category 2 to category 1 is assumed %
% to be estimated.                                                        %
%                                                                         %
% Notice that this script is easily extended to estimate more             %
% misclassification probabilities.                                        %
%                                                                         %
% Author: J.L.M. Bonekamp                                                 %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% output:                                                                 %
% nll     negative sum of the log-likelihood                              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% input:                                                                  %
```

```
% pars     coefficient for estimation first element is assumed to be alpha %
% I        element of matrix equals 1 when individual (row) is observed in %
%          that category (column) and zero otherwise.                      %
% X        matrix of covariates                                            %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[m, n] = size(X);

alpha = zeros(6,1);
alpha(4) = pars(3);

ALPHA = [1-alpha(3)-alpha(5), alpha(1), alpha(2);
         alpha(3), 1-alpha(1)-alpha(6), alpha(4);
         alpha(5), alpha(6), 1-alpha(2)-alpha(4)];

pars(1,:) = [];
pars = reshape(pars,n,2);

c_prob = ones(m,3);% probabilities to simply writing, comparable with p
for i = 2:3
    c_prob(:,i) = exp(X*pars(:,i-1));
end
sum_c_prob = sum(c_prob,2);
c_prob = c_prob./(sum_c_prob(:,ones(1,3)));

o_prob = (ALPHA*c_prob')';% observed probabilities, comparable with p^tilde

%increase efficiency with constraint optimization
G = sum(I.*o_prob,2)

Flag = ~(G~=G); G(~Flag) = 1e+308;

Lm=1e-323; G(G<=Lm) = Lm;
Hm=1-Lm; G(G>=Hm) = Hm;

l = log(G);
nll = -sum(l)/sum(Flag); % negative average log-likelihood
```