ISER Working Paper Series

Poverty persistence among Belgian elderly: true or spurious?

Marjan Maes
Catholic University Louvain
Catholic University Brussels

No. 2008-24
August 2008

INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

www.iser.essex.ac.uk

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Non-technical summary

In Belgium, the majority of people aged over 50 who are poor remain poor for a long time – they are persistently poor. There are a number of explanations for the persistence of poverty. On the one hand, poor people may have characteristics that make them particularly prone to poverty. These characteristics might be ones observed by researchers (e.g. low educational qualifications) or unobserved in a particular data set or be intrinsically observable (e.g. 'ability' of various kinds). On the other hand, it may be that poor individuals remain poor because the experience of poverty itself lowers the chance of escaping from poverty. People may become trapped in poverty because of e.g. a loss of motivation, or because employers use their poverty status as a signal of poor skills. Ascertaining the importance of these different explanations for poverty persistence is relevant for the development of anti-poverty programmes. If differences in characteristics play the major role, it suggests that measures should focus on improving skills and other personal characteristics. But if poverty itself plays a separate and independent role, there is a role for general policies aimed at improving structural features of the economy, e.g. removing adverse work incentives in the benefit system, or developing measures that reduce the 'scarring' effects of being poor.

This paper investigates the lengths of spells of poverty among Belgian people aged over 50 using statistical models that examine the determinants of the chances of leaving poverty among those who begin a poverty spell, and the chances of poverty re-entry among those who end a poverty spell. These models account for differences in personal characteristics (observed and unobserved). They also allow for the possibility that the chances of leaving poverty in any given year vary with how many years the person has already been poor. The existence of such 'duration dependence' provides prima facie evidence that poverty itself plays a role in determining persistent poverty.

The estimates of the statistical model lend support for the hypothesis that there is duration dependence in poverty in Belgium. This evidence is consistent with what is known about the Belgian economy. In particular, in Belgium, elderly unemployed people are not required to search for a job. This raises the chances that their skills will depreciate and that employers would be reluctant to invest in providing new skills or updating new ones for this group. In addition, both employers and the government provide pathways to retirement giving elderly people strong incentives to leave the labour market as soon as possible.

Poverty persistence among Belgian elderly: true or spurious?

Marjan Maes[*]

Abstract

Based upon a longitudinal administrative dataset merged with the Socio-economic Survey of 2001 and the National Register, the majority of the poor elderly in Belgium appear to be persistently poor. The simultaneous estimation of a multiple, spell discrete, time hazard model shows that dependence in poverty is a true phenomenon. It also shows that besides observable characteristics that reduce poverty exit and increase re-entry there are, in addition, unobserved effects that lead to the same kind of poverty persistence. Controlling for unobserved effects and an initial condition problem significantly improved the fit of the model.

Keywords: poverty dynamics; poverty persistence; early retirement; retirement incentives; multiple spell discrete-time hazard model

JEL codes: J14; J26; C41 ; I32

Correspondence to marjan.maes@student.uclouvain.be or marjan.maes@kubrussel.ac.be.
The author thanks Prof. Bart Cockx, Prof. Stephen Jenkins and participants to the workshop on social exclusion in Barcelona, the doctoral workshop 2008 in Louvain, ESPENET2008 in London and JESS seminar in Essex for critical comments and useful suggestions.

1. Introduction

OECD studies[1] and other studies[2] have argued that working is an effective means of staying out of poverty. At a first glance this seems difficult to reconcile with the fact that most OECD countries have designed social security systems that strongly encourage the elderly to stop working as soon as possible[3]. Given the fact that youth unemployment is left unaffected by early retirement of older workers[4] and that early retirement threatens the financial sustainability of social security systems, it becomes hard to understand why the Belgian government continues to encourage early retirement, especially if this would increase the risk of poverty among those who retire early.

In this paper I try to verify whether there is empirical support for the idea that adverse work incentives might push elderly into a poverty trap. This means that people may be given an incentive not to work while at the same time they slip into poverty.

One does not deny that early retirement may be the result of unfavourable labour demand and be perceived as an involuntary, forced[5], choice by the individual, "an offer he cannot refuse". On the contrary. But this does not impede the fact that the individual may at the same time have an incentive not to look for a job or become demotivated and discouraged. The question to be analysed is whether duration dependence in poverty among the elderly who are retiring from the labour force is true or spurious. To the extent that individual characteristics like low abilities or unfavourable attitudes persist over time, they may also be the reason that individuals persist in poverty over time. In that case, one expects that, once individual characteristics are controlled for, duration

[1] OECD (2005a), OECD (2005b), OECD Jobs Strategy (1994), OECD (1997), Förster (2000), Casey, Yamada (2002).

[2] On the basis of EU, SILC (2004, 2005), Zaidi, Makovec, Fuchs, Lipszyc, Lelkes, Rummel, Marin and DeVos (2006): "We also find that a large proportion of elderly have a high risk of persistent poverty. This can be true by default, since the elderly have little opportunities to enhance their income position in post, retirement life. Thus, the most effective policy intervention to enhance incomes of the elderly will be to increase incentives to work." See also Bardasi, Jenkins and Rigg (2000) for the UK.

[3] Fenge and Pestieau (2005); Blöndal and Scarpetta (1998); Grüber and Wise (1999); Grüber and Wise (2004).

[4] Boldrin, Dolado, Jimeno and Peracchi (1999); Fenge and Pestieau (2005); Jousten, Lefebvre, Perelman and Pestieau (2008).

[5] Dorn and Souza (2005); Lindeboom (1998); Jousten (2005)

dependence becomes spurious. The alternative possibility is that experience of poverty has a genuine causal impact on future poverty. According to Biewen (2003, 2004), this may be because low income may be associated with adverse incentives which make it not worthwhile for the individual to take up a job if unemployed or even to keep a low, paid job (the poverty trap). Or, this may be due to loss of motivation or depreciation of human capital which may lead to a series of low quality jobs or unstable employment, increasing the risk of re-entering poverty. The reasons for dependence in poverty are obviously of interest for developing effective poverty-reducing measures since true dependence would suggest a focus on stigma and adverse labour market incentives while spurious dependence would suggest a need to change individual's characteristics and abilities. Following Meghir and Whitehouse (1997), Devicienti (2002), Biewen (2003), Fertig and Tamm (2007) and Hansen and Wahlberg (2004) the question whether duration dependence reflects true duration dependence or individual heterogeneity is analyzed through a multiple-spell model of transitions in and out of poverty, controlling for observed and unobserved individual heterogeneity and for a potential initial condition problem. For Belgium, this issue has been addressed for welfare spells (in Cockx, 1998) and for unemployment spells (Cockx and Dejemeppe, 2005).

This paper is structured as follows: Section 2 explains the matching between the administrative dataset and the Socio-economic Survey and how individuals that are temporarily out of the Income Tax Returns have been integrated. Although this might seem an innocuous operation, it is not because it implies choices to be made on the concept of income, household unit and equivalence scales. Section 3 describes at the aggregate, population, level the observed poverty transition rates associated with the duration of a (non)poverty spell. If there are individual-specific unobserved factors that affect the hazard, the aggregate transition rates will tend to be different from those at the individual level. In order to examine this, section 4 estimates a multiple-spell discrete-time hazard model through which I simultaneously estimate exit and re-entry rates while allowing for observed and unobserved heterogeneity and controlling for a potential initial condition problem. This section also contains an overview of the modelling literature on poverty dynamics. Section 5 concludes.

2. Data Construction

The results presented below are derived from a micro-dataset provided by the National Institute for Statistics that contains information on 30,183 Belgian private[6] households[7] of which at least one member was between 55 and 75 years old on 31 December 2001 and that were randomly selected from the National Register. Of these 30,183 households, one keeps the individuals that are between 55-75 years old in 2001. This reduces the number of individuals of the final dataset from 60,806 to 43,726.

These 30,183 households (corresponding to 43,726 individuals) have been linked to the Income Tax Returns (ITR) data (1991 – 2002)[8] by means of the national identification number. This implies that the individuals in our dataset are between 45-65 years old in 1992, 46-66 in 1993, 47-67 in 1994 and so on. This administrative dataset contains the annual information necessary to calculate the income tax. The variables it includes are civil status, number and type of dependants in the fiscal household, gross capital income, the age and gross labour income of the household members, replacement incomes of the household members (old-age pension, early retirement, unemployment, illness or disability benefits etc), housing wealth, pension benefits from the second and third pillar and employee contributions in the second and third pension pillars. Every year, about

[6] Collective households are excluded from the population subject to sampling.

[7] That is 60,806 individuals. A household is defined as the number of individuals having the same domicile, as registered in the National Register. On the basis of the National Register of 2001, 93.66% of the 60,806 individuals are head (49,6%), partner of the head (29.35%) or child (14.72%) of the head of the household. Only 6.33% has another kind of relationship to the head: 1.25% are grandchild, 0.76% parent of the head, 0.47% child, in, law, 0.48% parent in law, 0.46% brother or sister and 2.22% are no family related habitants.

[8] Note that 723 of the 30,183 households group several fiscal household units. For example, two widowers/divorced/single individuals living together but filling in separate tax files are two fiscal households but will be considered as one household. They have similar socio-economic characteristics as the other households except that they contain 3 times more single and legally divorced individuals, 10 times more effectively divorced individuals, many more ‘no family related’ habitants, 10% more females and 10% more individuals younger than 65. However, the fact that we dispose only of data of the National Register for 2001 may generate measurement errors concerning the composition of such households before 2001. One does not know whether before 2001 these widowers or divorced individuals were living in the same house. First, we suppose they are living together from the first year in which they are observed both as declarant in the ITR data while it may be they start living together only a few years afterwards. Secondly, it may be that of the two persons living together in 2001 one or both have lived with somebody else before 2001 with whom they formed a fiscal household. This concerns only 0.03% of the households. For these years before which both persons become declarant, we only retain as household the couple that contains the individual that will be head in 2001 and drop the observations of the other fiscal household.

86% of the individuals selected from the National Register could be matched with the ITR. This means that in a given year about 14% of the Belgian civil population are not in the ITR.[9] However, only 4.2% of the individuals selected from the National Register and between 55-75 years old in 2001 (that is 1,844 of 43,726) or 4.9% of the households (that is 1,502 of 30,183), do not appear in the ITR for any year 1991-2002 and will have to be deleted from the analysis altogether. We will discuss this issue below. The stability over time in the percentage of individuals out of the ITR thus masks considerable income mobility at the individual level in and out of the ITR.

The 30,183 households (corresponding to 43,726 individuals) selected from the National Register could also be merged with the Socio-economic Survey of 2001 through the use of the national identification number. This survey has a response rate of 98.7% at the individual level and 98.6% at the household[10] level. It contains detailed information on education level, professional category (private sector employee, civil servant, self, employed etc), the sector in which the household member works or worked (agriculture, banking, insurance, construction, transport, chemical industry, real estate, army, education, retail etc) and also for the first time in Belgium the self-reported general health status.

[9] This confirms the finding of Perelman, Schleiper and Stevart (1998) that 13% of the Belgian population do not declare incomes.
[10] We consider a household as participating if at least one household member participates.

Table 1: Construction Dataset		
	Number of households	Number of individuals
National Register	30183	60806
National register 55-75 years old	30183	43726
National Register 55-75 years old,	30183	43726
matched with Socio-economic Survey of 2001	29760 (matching: 98,6%)	43157 (matching: 98,7%)
matched with ITR		
1991	25830 (matching:86,3%)	37423 (matching:85,7%)
1992	25515 (matching: 84,5%)	37023 (matching: 84,67%)
1993	25559 (matching: 84,6%)	37115 (matching: 84,88%)
1994	25326 (matching: 83,9%)	36856 (matching: 84,29%)
1995	24778 (matching: 82,0%)	36208 (matching: 82,80%)
1996	25566 (matching: 84,7%)	37213 (matching: 85,10%)
1997	25727 (matching: 85,2%)	37465 (matching: 85,68%)
1998	25850 (matching: 85,6%)	37651 (matching: 86,10%)
1999	25855 (matching: 85,6%)	37677 (matching: 86,16%)
2000	25914 (matching: 85,8%)	37764 (matching: 86,36%)
2001	25982 (matching: 86,0%)	37819 (matching: 86,49%)
2002	26269 (matching: 87,0%)	38213 (matching: 87,39%)
Number of individuals (households) that never appear in ITR between 1991-2002: 1,844 (1,502[11])		
Number of individuals (households) that appear at least one year in ITR between 1991-2002: 41,882 (28,681[12])		

As can be seen from table 1, the number of households is not the same for every year of the ITR: households may temporally or permanently drop out of the ITR. According to

[11] 1635 households with a member that does not declare , 133 households where at least one member declares income.

[12] 28548 households where every member declares+133 households where at least one member declares income.

article 178 of the Royal Decree corresponding to the Belgian income tax code of 1992, those not obliged to declare incomes are: 1° households without professional activity with an income below the minimum taxable income (except singles/widow(er)s with dependent children) and 2° households of which the income only consists of old-age pensions and housing wealth. It is unfortunately[13] impossible to know whether a dropout is due to 1° or 2°[14]. Individuals that drop out because of reason 1° may, from the moment their income exceeds the minimum taxable income, reappear in the ITR. If that happens we qualify missing periods before the reappearance as periods in poverty. Eligibility rules for old-age pension benefits can also be used to qualify missing observations as periods in poverty. Table 2 shows how this correction reduces the number of unbalanced households. Individuals that drop out because of reason 2° or because of reason 1° that do not reappear will, from that moment on, never appear again in the ITR data. We treat these drop-outs as right-censored observations[15].

[13] We contacted the fiscal administration who affirmed that tax officers in some tax localities themselves add a code for 1° or 2° in the tax files but this is not systematically done and thus not reliable information. Another technical possibility was to merge the data with an existing dataset of means-tested beneficiaries for the years 1991-2001 but access to this dataset was not allowed, although this would enable us to identify (part of) 1°.

[14] A dropout cannot be due to death since the sampling scheme implies that individuals are sampled conditional on being alive in 2001. If poor singles would be more likely to die this could induce an endogenous selection problem. Up to this stage, this has been neglected.

[15] Following the procedure used by Devicienti (2002), Stevens (1999), Devicienti and Gualtieri (2007), Fertig and Tamm (2007). However to the extent that the dropouts because of reason 1° cannot be considered as random censoring, the sample selection problem they might induce should be modeled explicitly. This is an issue for future research. Capellari and Jenkins (2004) model attrition simultaneously with poverty transitions for UK and Lillard and Panis (1998) model household composition and attrition simultaneously for the US but both find that attrition induces a negligible bias in the estimation results.

Table 2: Unbalanced data				
Number of years observed	Number households after correction	% after correction	Number households before correction	% before correction
1	279	0.97	671	2.34
2	199	0.69	469	1.64
3	217	0.76	499	1.74
4	229	0.80	476	1.66
5	277	0.97	567	1.98
6	266	0.93	548	1.91
7	359	1.25	651	2.27
8	380	1.32	689	2.40
9	481	1.68	825	2.88
10	453	1.58	1169	4.08
11	845	2.95	2743	9.56
12	24696	86.11	19375	67.55
TOTAL	*28681*	*100.00*	*28681*	*100.00*

The correction reduces the number of missing observations of individuals that appear at least once in the ITR from 11.9% to 5.4% of total observations[16]. The effect of this correction is to increase poverty persistence, reduce transitory poverty, increase the number of households that is confronted with poverty once and to change the coefficients of duration dependence in the regressions. However it does not change any of the general conclusions on the issues under study.

The raw income data of the final dataset are adjusted in three steps to give more information about the well-being of the households. First, since net income is a better indicator of the living standard of the household than gross income, we calculate for every household its net income. Secondly, since all income data are nominal and the data is a time-series from 1992 to 2002, they were converted into real data[17] in order to represent purchasing power of households. In a third step we make net incomes

[16] The number of missing observations (41,190 out of a total of 344,172 observations (12 years for 28,681 households that appear at least once in the Income Tax Returns data)) could be reduced to 18,888 observations. We also qualified the missing observations (1,792; 4.3% of 41,190) of households with 0 years in poverty, at least 10 years out of poverty and mean income pooled over all available years of more than 140% of the poverty line as non-poor. This allows the incorporation of missing observations that correspond to measurement errors of old-age pensioners or temporary emigration but not to a poverty experience.

[17] With year 2002 as reference year for all individuals.

comparable between households of different sizes and with different needs[18] by inflating net real income by the OECD equivalence scale that attributes 1 to the head of the household, 0.5 per additional adult and 0.3 per child.

In the remainder of this paper the focus of attention will be the head of the household[19] but the unit for calculating income is the household. This means the head of the household is qualified poor if the income of the household to which he pertains is below the poverty line. The latter is defined as 50% of median net equivalised income[20] of the whole economy.

3. Descriptive analysis of the dynamics of poverty[21]

The longitudinal dataset presented in the previous section is well-suited to describe the dynamics of poverty. On the basis of the dataset, I will start with a description of flows into and out of poverty and the distribution of periods spent in poverty to give an idea of the extent to which Belgian elderly that become poor persist in poverty.

3.1. To start, table 3 shows the distribution of the total number of years spent in poverty.

---

[18] For two widowers or divorced living together, net incomes are calculated separately for each fiscal household. Then the net income of all fiscal households is summed and in a last step the equivalence scale is applied on this sum.

[19] The head of the household is the individual that declares income. For married individuals, the fiscal legislation says it is the man. In 2,838 households the spouse of the declarant becomes declarant herself due to death of her husband or divorce, so that in the year 2001 of the survey the husband is no more part of the household. In that case, when we decompose FGT indices by socio-economic characteristic of the head, the latter correspond to those of the women. Similarly, if the head of the fiscal household is not 55-75 years old in 2001, but the partner of the head is, we take the partner as head of the household (1,983 households). In cases with two fiscal households living together (723 households),the head is defined as the member that is most years in the ITR.

[20] That were kindly provided by the National Institute for Statistics, for the years 1992-2002 and for different equivalence scales. However equivalence scales taking into account the number of disabled in a household are not at our disposal.

[21] Excluding the ,1844 individuals that are in the Socio-Economic Survey but not in the ITR.

Table 3: distribution of total number of years spent in poverty[22]			
Number of years in poverty	Number of households	% of those who are poor at least once	% of population
0	17997	/	0.63
1	2435	0.22	0.085
2	1212	0.11	0.042
3	944	0.086	0.032
4	721	0.066	0.025
5	781	0.071	0.027
6	624	0.057	0.021
7	586	0.053	0.020
8	513	0.047	0.017
9	548	0.050	0.019
10	577	0.053	0.020
11	644	0.059	0.022
12 or more	1099	0.101	0.038
Number of households that are at least 1 year in poverty: 10,684			
Number of households: 28,681			

The fact that 63% of the elderly households are never poor implies that 37% of them are confronted with poverty at least once. This is much higher than the "static" poverty rates of around 12% that are found on the basis of the PSBH in Cantillon (1999) and also on the basis of our own dataset. Deleeck and Cantillon (1992) find similarly on the basis of two waves of the SEP that of the whole population 10.8% are poor[23] in 1988 and 1985, 73% are not poor in 1988 nor in 1985 while 16.2% are poor once during that period.

For our data the transitory poor, who are poor for one year, account for 8.5% of all households. Those who are poor for at least 3 years make up 66% of those who ever have

[22] Including left-censored spells
[23] Poverty line is 50% of average income.

been poor and 24% of all households. In general most people that slip into poverty are quite successful in getting out. But precisely because this is true the transitory poor are a small fraction of the poor at any point in time and those with longer poverty spells account for the bulk of all poverty.

3.2. The degree of persistence of poverty and the recurrent poor cannot be read from table 3. Households who are poor for three years are not necessarily in poverty for three *consecutive* years. The *persistently* poor are poor for at least three consecutive years. The *recurrent* poor are poor for at least one year but never longer than two consecutive years. It is thus possible that a recurrent poor household is, for example, poor for five years in total but is not persistently poor. Table 4 shows that more than 60% of the elderly who once have been poor are persistently poor.

Table 4: Persistent and recurrent poverty among elderly[24]

	Number of households	% of population	% of those who are poor at least once
Persistent	6499	0.22	0.61
Recurrent poor	4185	0.14	0.39
Number of households that are at least one year in poverty: 10,684			
Number of households: 28,681			

3.3. The number of consecutive years one is in poverty defines a poverty spell. When studying poverty spells the issue of censored spells arises. Suppose a household is counted poor for exactly one year. We would qualify it as transitory poor. However if that year corresponds to the first/last year of observation, the duration of the poverty spell is underestimated if the household was poor before the sampling began/after the sampling stopped. The following table displays the considerable proportion of censored poverty spells. It will be discussed below how censored spells will be taken account of in the multivariate analysis.

[24] Including left-censored spells.

Table 5: Censored poverty spells			
	Number of households	% of population	% of those who are poor at least once
Left-censored	5082	0.17	0.47
Right-censored	5200	0.18	0.48
Left- and right-censored	2758	0.09	0.25
Number of households that are poor at least once: 10,684			
Number of households: 28,681			

3.4. From the moment that the data are arranged such that each household is associated with the duration of one (or more) spells, one can calculate exit and (re)entry rates. The exit rate associated with a given duration of a spell is the number of households that exit at that length of the spell divided by the population at risk of exiting. The survivor function associated with a certain spell duration specifies the probability that an individual will survive in that spell beyond that duration. Spells that are right-censored are included in all but the censored year. Figures 1 and 2 plot the poverty exit rates and survivor function with and without left-censored spells but they do not differ a lot from each other[25].

[25] The exclusion of left-censored spells implies that exit rates can only be calculated for a maximum duration of 10 years.

Figure 1: Survivor function with and without left-censored spells

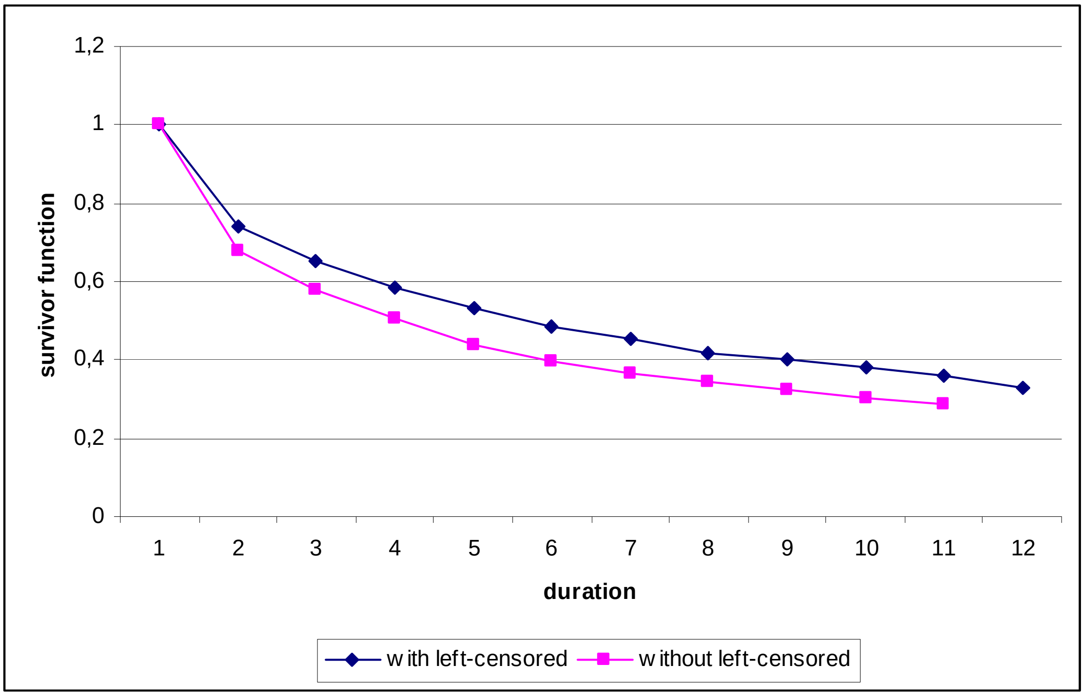
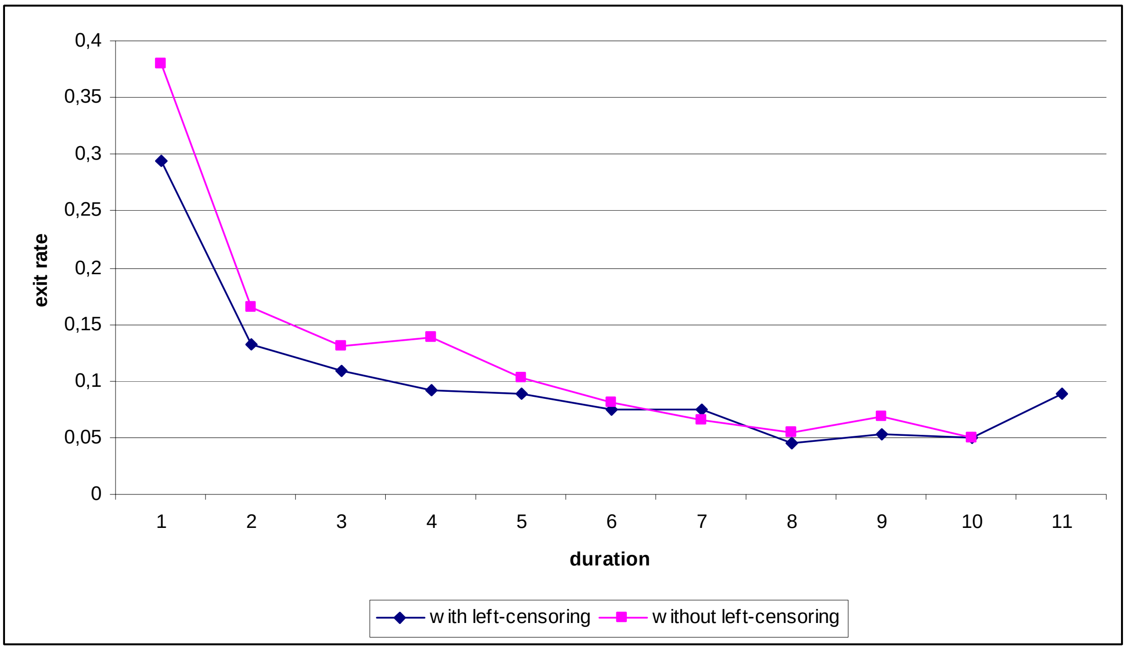


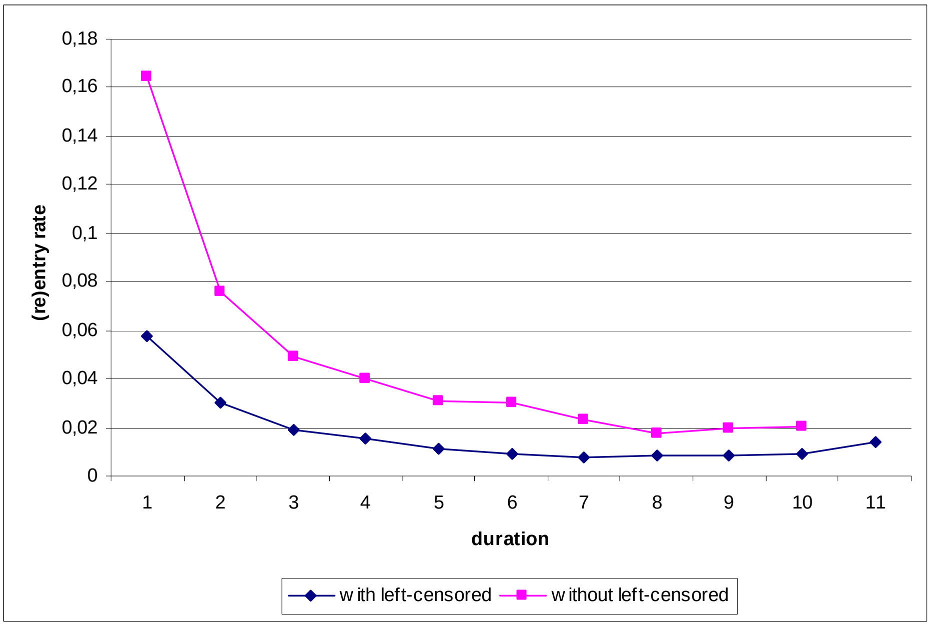
Figure 2: Exit probability with and without left-censored spells



The exit rate is high for the transitory poor and low for those that are long-term poor. The survivor function decreases sharply in the first years but after some years seems to stabilize. This means that the longer somebody is poor, the more difficult it becomes to leave poverty.

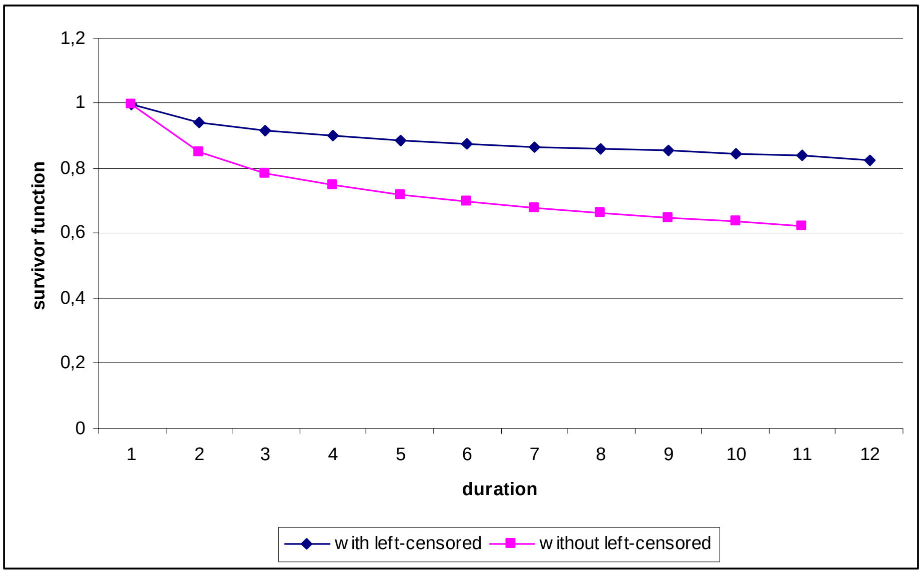
To find the (re)entry probabilities, one calculates for each of the households at risk of (re)entering poverty the length of the spell that they are out of poverty. Then for each possible length of the spell the number of individuals that enter poverty is divided by the population at risk of (re)entering. The results, whether left-censored spells are included or not, are in figures 3 and 4 and differ quite a lot[26]. The re-entry rates with left-censored spells are commonly called entry rates. Probabilities of entering poverty are very low at around 1-2%: they are based on households that may or may not have been poor once. Re-entry rates are up to 4-6 times higher than the entry rates, indicating that the probability of becoming poor is much higher for households who have been poor than for those who have not.

Figure 3: (Re)entry probability with and without left-censored spells



[26] The exclusion of left-censored spells implies that the entry rates can only be calculated for a duration of maximum 10 years.

Figure 4: Survivor function with and without left-censored spells



Survivor functions represent the cumulative risk of slipping back into poverty after a previous exit. Of those who quit poverty, about 60% are still out of poverty after 12 years. Thus 40% of those who quit poverty will fall back into poverty.

The question that interests us is whether the exit rates out of poverty and re-entry rates into poverty, that are downward sloping, remain so after one has controlled for all kinds of individual effects. To the extent that individual characteristics, like the lack of abilities, persist over time they may be the reason that the Belgian elderly persist in poverty over time. If this would be true, one expects that after controlling for individual-specific effects, duration dependence would no longer be significant. Alternatively, if current poverty experience has a causal impact on future poverty experience, one expects that negative duration dependence remains significant. To check this, a multivariate analysis is necessary, to which we turn now.

4. A multiple-spell discrete-time hazard model of poverty dynamics

The purpose of this section is to verify whether duration dependence in the exit and re-entry rates is due to individual heterogeneity or true duration dependence through the estimation of a multiple-spell discrete-time hazard model while controlling for unobserved effects and a potential initial condition problem (4.3.). We start this section with an overview of the existing empirical models of poverty dynamics (4.1.), followed by a presentation of the model that will be used for our purposes (4.2.).

4.1. Previous modelling research on poverty dynamics

This section briefly reviews the empirical literature because to demonstrate why some type of model are not used and why only one type of model is appropriate for my purpose. In the end, I discuss the few models that have been estimated on Belgian data.

4.1.1. Component of variance models

One of the first papers to study poverty dynamics was Lillard and Willis (1978) who estimate an earnings model with log of earnings y_{it} as the dependent variable, individual $i = 1,...,N$ at time $t = 1,...,T$ and x_{it} a vector of observed explanatory variables:

$$y_{it} = x_{it}\beta + v_{it}$$

The error structure has the form: $v_{it} = c_i + \gamma v_{it-1} + u_{it}$ with c_i an unobserved effect, u_{it} a random error term and γ a serial correlation coefficient common to all individuals. However, if the serial correlation in the error structure results from misspecification of the population model: $y_{it} = x_{it}\beta + \rho y_{it-1} + c_i + u_{it}$, c_i is correlated with y_{it-1} and the use of instrumental variables may be necessary. Up to now, these models did not address either the fact that explanatory variables such as household composition or labour market status that are often included might be endogenous to the dynamics of income. In addition, a common variance structure is assumed for the entire population while the dynamics of individuals in different parts of the income distribution might be different. According to Stevens (1999) and Devicienti (2002), they perform less than duration models in predicting poverty. A recent application is Fouarge and Muffels (2003).

4.1.2. Duration models and transition probability models

Bane and Ellwood (1986)[27] calculate duration-dependent exit probabilities and the distribution of entering poverty and of exiting poverty as a function of events (change in household composition or household income). They also calculate the expected duration of poverty as a function of events associated with the beginning of a poverty spell. Although this is not a multivariate analysis, they initiated a new strand of a literature that analysed the determinants of flows into and out of poverty.

Transition probability models are developed with probability of (re)entry in and exit out of poverty as the dependent variable and change in employment status and/or household composition, components of household income, individual and household characteristics as independent variables [28]. At the same time models arise that estimate the probability of (re)entry in and exit out of poverty as dependent variable with dummies for duration, individual and household characteristics as independent variables. The latter are in fact equivalent to discrete-time hazard models, as shown by Allison (1984) and Jenkins (1995).

More complicated discrete, time hazard models extend the analysis from single to multiple spells. Most of these estimate poverty transition equations separately[29] under the

[27] For the US. For the UK, see Jenkins (1998), Jenkins (2000), Jenkins and Rigg (2000). For Germany, see Otto and Siedler (2003). For Germany, US, UK, Canada, see Antolin, Dang and Oxley (1999).
[28] Bourreau, Dubois, Jeandidier and Berger (2003); Bardasi, Jenkins and Rigg (2000); Zoyem (2002); Valletta (2005); Dewilde (2004); McKernan and Ratcliffe (2002); Antolin, Dang and Oxley (1999); Fouarge and Layte (2003).
[29] McKernan and Ratcliffe (2002); Finnie and Sweetman (2002); Devicienti (2001a); Antolin, Oxley and Dang (1999); Zoyem (2002); Fouarge and Layte (2003); Valletta (2005); Dewilde (2004); Makovec (2005); Bourreau, Dubois, Jeandidier and Berger (2003); Devicienti (2007); Capellari (2007); Jenkins and Rigg (2001); Arranz and Cantto (2006). Arranz and Cantto (2006); McKernan and Ratcliffe (2002); Finnie and Sweetman (2002); Antolin, Dang and Oxley (1999); Fouarge and Layte (2003) and Devicienti (2001a)

hypothesis that, for a given individual, entry rates and exit rates can be treated as conditionally independent and that multiple spells of the same event type are conditionally independent. A first extension to this, as Callens, Croux and Avramov (2003) and Arranz and Cantto (2006) show, is to allow for different baseline hazards in case of multiple spells of the same event type for a given individual. An additional extension captures not only correlation between spells of the same event type but in addition that individuals with high (low) exit rates have lower (higher) re-entry rates. Stevens (1999)[30] is the first to estimate exit and re-entry equations of poverty simultaneously while allowing the unobserved effect of the two transition equations to be correlated. This leads to more accurate estimates of total time spent in poverty. The same method of joint estimation of entry and exit equations is applied by Jenkins and Rigg (2001), Devicienti (2002), Biewen (2003), Wahlberg and Hansen (2004), Fertig and Tamm (2007) and Devicienti and Gualtieri (2007). For these joint estimations, the most frequent distributions for the unobserved effect are the discrete distribution with a finite number of support points [31] and the multivariate normal distribution[32].

There exist also discrete-time duration models that estimate exit and re-entry rates separately while accounting for unobserved heterogeneity: Finnie and Ross (2002), Fouarge and Layte (2003), Makovec (2005) and Capellari (2007). However, it appears[33] that for single-spell data, estimation results are sensitive to misspecification of the distribution of unobserved effects while for multiple-spell models, it is much easier to estimate parameters that are robust to the functional form of the unobserved effect[34].

---

include however the number or length of previous spells, although this is an endogenous variable. Note also that even if the exit and entry processes would be independent, estimating them jointly instead of separately would be more efficient.

[30] Following Meghir and Whitehouse (1997) who estimate jointly unemployment and employment spells, while accounting for unobserved effects and an initial conditions problem.

[31] Meghir and Whitehouse (1997); Stevens (1999); Devicienti (2002); Hansen and Wahlberg (2004); Biewen (2003); Fertig and Tamm (2007) use the Heckman and Singer (1984) estimator.

[32] Lillard and Panis (1998).

[33] Heckman and Singer (1984). Meyer (1990, p.771) notes that "it is plausible that much of the parameter instability found by Heckman, Singer (1984) is due to their assumption of a Weibull baseline hazard. When the baseline hazard is nonparametrically estimated, the choice of heterogeneity distribution may be unimportant". Nicoletti (2006,p.19) finds in the same line for discrete, time hazard models that "misspecifying the random effects distribution biases neither the duration dependence nor the covariate coefficients estimation".

[34] van den Berg (2001).

The discrete-time hazard model takes account of right-censored spells[35] under the assumption they are randomly censored but left-censored spells are more problematic. Excluding them, as a lot of models do[36], could result in a sample selection bias in the presence of unobserved heterogeneity. Lancaster (1990) notes: “The common treatment of stock sampled data with future spells observed is to ignore the elapsed duration data and to base inferences solely on those spells that begin after the sampling data. This is a sensible and correct way to proceed in models that do not involve unmeasured person, specific heterogeneity. Unfortunately, in models that do involve such heterogeneity there is a further complication to consider due to the fact that the distribution of unobservable quantities also depends of the sampling scheme”[37]. If the probability that the first spell will be poverty or non-poverty depends on individual characteristics, including any unobserved heterogeneity, excluding left-censored spells creates an initial conditions problem whereby the identity of the first spell is endogenous. To control for this selection bias, Devicienti (2002), Biewen (2003), Wahlberg and Hansen (2004) and Fertig and Tamm (2007) extend the analysis of Stevens (1999) and estimate the entry and exit equations jointly with an initial condition equation, as Meghir and Whitehouse (1997), while allowing the unobserved effects in the transition equations to be correlated with the unobserved effect in the initial condition equation. In order to identify the model, the initial condition equation contains exclusion restrictions. Whitehouse and Meghir (1997) use unemployment rates at the first time the spell is observed, Fertig and Tamm (2007) use dummies for the year of the first observation and education level of the parents of the household head, Biewen (2003) uses the education level of the parents and city where the individual grew up, Devicienti (2002) uses the education level of the parents of the household and Hansen and Wahlberg (2004) use no exclusion restrictions at all which

[35]Each individual contributes to the likelihood function for every year he is in the dataset but depending on whether the last interval is censored or not the dependent variable of the last contribution equals 0 or 1. It is important to note that simply eliminating all the censored observations leads to a different likelihood function and might bias the estimation results.
[36] Bane and Elwood (1986); Antolin, Dang and Oxley (1999); Finnie and Sweetman (2002); Devicienti (2001a , 2002); Fouarge and Layte (2003); Makovec (2005); Devicienti (2007); Arranz and Cantto (2006). The problem of left, censoring is absent in the case of flow sampling (Cockx, 1998; Dejemeppe, Cockx, 2005).
[37] P.189.

implies that the model is supposed to be identified by the functional form of the unobserved effects. Another possibility to treat the sample selection bias induced by stock-sampling is to include left-censored spells but to correct the likelihood function for the fact that the length of the first spells is underestimated. This requires the assumption of a constant entry rate and a constant survivor function, or additional data such as the aggregate number entering poverty at each calendar date in the past as is empirically relevant such that one can construct, as Nickell (1979), a model for the entry rate and survival function. There are no such data for Belgium and thus one cannot apply this method.

4.1.3. Dynamic unobserved effect models

Dynamic unobserved effect models are developed in unemployment dynamics literature, by Heckman (1981), Arulampalam, Booth and Taylor (2000), Stewart and Sheffield (1999), Stewart (2007), and also have found poverty applications. Specify the model for individual i=1,…,N at time t=2,…,T as

$$y_{it}^{*} = x_{it}^{'}\beta + \gamma y_{it-1} + c_i + u_{it}$$

Where y_{it}^{*} denotes the unobservable propensity to be poor, y_{it-1} the observed poverty status in t, 1, x_{it} a vector of observable characteristics, c_i an unobserved effect and u_{it} a random error term. y_{it} is the dependent variable

$y_{it}\begin{cases} =1 & if \quad y_{it}^{*} > 0 \\ =0 & else \end{cases}$. The inclusion of the lagged dependent variable allows to test for state dependence. In contrast to discrete-time hazard and components of variance models that take into account a history of lags, this model assumes thus that one lag of poverty status is sufficient to capture the full dynamics. According to Devicienti and Gualtieri (2007) the significant duration dependence in observed transition rates casts doubts on the first-order lag often assumed in empirical work. To estimate this model, one first integrates out the unobserved effect by assuming a distribution for c_i that is independent from x_{it}, usually a normal or Gamma. Secondly if the initial observation of y_{it} is correlated with c_i, this raises the initial conditions problem. Accordingly, Heckman (1981) and Cappellari, Dorsett and Haile (2007) estimate an initial condition equation jointly with the poverty transition equation while allowing for correlation between the unobserved effects affecting the poverty transition and initial condition equations. With the idea that initial labour market states, but not transitions, depend upon the macroeconomic conditions prevailing at the time, they use GDP growth rate measured in the year of the first observation as an exclusion restriction. Another approach has been suggested by Wooldridge (2002) that consists of modelling the distribution of the unobserved effect conditional on the initial poverty status. It is applied by Poggi (2007) and Biewen (2004)[38]. In addition, Biewen (2004) challenges the assumption of strict exogeneity of the explanatory variables, that is that there must not be any feedback from current poverty to future values of the explanatory variables; employment status and household composition. To this end, he estimates a joint dynamic random effects model of poverty status, employment status and household composition status. Since the assumption of one lag of poverty status is unreasonable in our case, this model was not estimated.

4.1.4. Markovian transition models

Cappellari (1999), Capellari and Jenkins (2002), Capellari and Jenkins (2004) and VanKerm (2004) model entry and exit probabilities simultaneously using an endogenous switching regression model with a binary dependent variable[39].

Specify the poverty transition equation for individual i= 1,…,N at time t = 2,…,T as

[38] McKernan and Ratcliffe (2002) condition the unobserved effect on the initial poverty status in a discrete-time hazard model. They find that individuals with left-censored (non)poverty spells are significantly less likely to (enter)exit poverty.
[39] Following Stewart and Swaffield (1999) who used this model to estimate low pay dynamics.

$$y_{it}^* = \left[(y_{it-1})\gamma_1' + (1 - y_{it-1})\gamma_2' \right] x_{it-1} + c_i + u_{it}$$

With y_{it}^* the unobserved propensity of being poor in t, y_{it-1} poverty status in t-1, γ_1' and γ_2' the coefficient estimates conditioning on being poor respectively non-poor in t-1, x_{it-1} a vector of observable characteristics, c_i an unobserved effect and u_{it} a random error term. y_{it} is the dependent variable

$$y_{it} \begin{cases} = 1 \quad if \quad y_{it}^* > 0 \\ = 0 \quad else \end{cases}$$

These models typically account for different sources of non-random selection such as attrition and initial conditions by estimating jointly an initial condition equation and an attrition equation with the poverty transition equation, while allowing the unobserved effects of these equations to be correlated with a multivariate normal distribution. Heckman (1981) suggested the use of pre-sample information such as the education level, occupation, labour market status of the parents of the household head as exclusion restrictions. Model estimates can be used to derive predictions of the poverty persistence rate and entry rate. In contrast to the dynamic unobserved effect models, the lag structure in x_{it-1} rules out the possibility of instantaneous effects of changes in characteristics for poverty status. For example, changes in employment status are not allowed to affect poverty until the next period. In addition, one may have doubts on the appropriateness of the first-order dynamics assumption[40]. That is why this model is not used.

4.1.5. Structural models

Aasve, Burgess, Dickson and Propper (2006) argue that poverty is not a decision variable but rather the outcome of underlying behavioural decisions such as whether to work, to have children, to marry and divorce. They estimate five simultaneous hazards (childbearing, marriage, divorce, employment and non-employment) while allowing the

[40] Devicienti and Gualtieri (2007).

unobserved effects of these five equations to be correlated according to a multivariate normal distribution. From these results, one can derive a model for income dynamics and poverty. In their model all persistence within poverty is attributed to persistence within demographic and labour market states. Although already very time-consuming and complex, they only account for two labour market statuses such as employed versus non-employed, while we are interested in the transitions between disability, unemployed, self-employed or employed, old-age pension and early retired. In addition, there is no consensus on whether these labour market transitions are voluntary or involuntary: early retirement may result from an unfavourable labour demand. Finally, changes in household composition, like becoming widowed, cannot easily be treated as a behavioural decision. For these reasons, this model was not estimated.

After this overview of the existing research on modelling poverty transitions, let us look what models have been applied to Belgian data. In Belgium, models on the duration of unemployment spells (Dejemeppe and Cockx, 2005) and welfare spells (Cockx, 1998) are available, as well as models on poverty persistence, but not for the elderly. Van Kerm (2004) finds, through the estimation of a Markov transition model on the PSBH on a population between 25 and 55 years old, that poverty entry depends on household and employment status. In particular being unemployed, self-employed or single increases the risk of poverty entry. He controls for the endogeneity of the initial poverty status by estimating jointly with the poverty transition equation an initial condition equation while allowing the unobserved effects of these equations to be correlated with a trivariate normal distribution. He does not consider the possibility that household and employment status that are included as explanatory variables might be endogenous. Indeed, current poverty status may affect future employment and household composition. Similarly, Nicaise and Deblander (2005) estimated a Markovian switching model on the PSBH for the years 1993-1997 for the working age population, while controlling for initial condition equations. As an extension to VanKerm (2004), Nicaise and Deblander (2005) controlled for possible endogeneity of employment status (but not for household status). They find that initial employment status is insignificant to explain initial poverty status and transitions into poverty but is significant in explaining poverty exit[41]. Household composition has a significant effect on poverty transitions. Dewilde (2004) finds through the separate estimation of a transition probability model on the PSBH for the whole population that poverty entry and exit depend on household and employment related events. In particular entry into unemployment, disability, (early) retirement is associated with entry into poverty. Makovec (2005) uses a discrete-time hazard model to estimate separately poverty entry and exit equations for those above 55 years old in the European Community Household Panel (ECHP), taking into account unobserved heterogeneity. For the entry model, he finds that accounting for unobserved effects leads to spurious dependence. Like Vankerm (2004), Makovec (2005) includes household and employment status as explanatory variables, ignoring possible endogeneity. He includes dummies for employed versus non-employed, for those aged above or below 65, for receipt of disability and old age benefits. None of the above dummies is significant except the receipt of disability benefits, that increases poverty exit. That is the opposite result of Dewilde (2004) although the dataset in the ECHP for Belgium is also the PSBH. Dewilde (2004), though, does not control for duration effects, nor for unobserved effects that are highly significant in Makovec (2005). On the other hand, Makovec (2005) used a very small sample limited to those above 55 years old while Dewilde (2004) also reports problems with limited sample size. Although he conditions on unobserved effects, Makovec (2005) does not address the issue of left-censored spells.

4.2. The model

Following Meghir and Whitehouse (1997), Stevens (1999), Devicienti (2002) and Biewen (2003), I estimate the discrete-time hazard model, given the interest in duration dependence. Although in the real world poverty transitions can occur at any time, the

[41] Instead of a simultaneous estimation of poverty and employment equation, they followed a less efficient two-step procedure. 1) One estimates an initial employment equation and saves the generalised residuals and an initial poverty equation where one saves the generalised residuals. 2) one inserts the residuals of the initial employment equation in the employment equation of interest that is estimated and one saves again the generalised residuals. 3) one inserts the generalised residuals of the initial poverty equation and the generalised residuals of the employment equation in the poverty transition equation of interest that is estimated. The exit and entry equations are estimated separately.

model is in discrete time since the data are grouped into intervals of one calendar year. There are two types of spells: poverty spells (k=p) and non-poverty spells (k=np). We assume that the probability that an individual i=1,…,N leaves the spell of type k in the calendar year t=1,..,T at a duration d is defined as $P_{it}^{k} = \Pr ob(y_{it}^{*k} > 0)$ that results from the latent model $y_{it}^{*k} = c_i^k + f^k(d) + \beta^k x_{it} + u_{it}^k$ where y_{it}^{*k} denotes the unobservable propensity to be in a spell of type k, x_{it} is a vector of observable characteristics, β^k the vector of coefficients associated with x_{it}, $f^k(d)$ is a function of duration dependence that represents the baseline hazard and where d=1,…,D denotes the duration of the current spell and D is the maximum duration of a spell. We will adopt a flexible specification for the baseline hazard[42] where $f^k(d) = \alpha_1^k DU_1 + \alpha_2^k DU_2 + ... + \alpha_D^k DU_D$ and DU_d are dummies corresponding to a duration d. We assume that individuals enter a spell at d=0 and are at risk of leaving the spell at d=1,….,D[43]. Unobserved heterogeneity enters the specification of the hazard rate as an individual-specific additive error term c_i^k constant over time which is allowed to be correlated across different types of spells. We are thus estimating a binary response model where the dependent variable

$y_{it}\begin{cases} =1 & if \quad y_{it}^{*} > 0 \\ =0 & else \end{cases}$ and u_{it}^k is a random error term. If u_{it}^k is logistically[44] distributed,

$P_{it}^{k} = \dfrac{\exp(c_i^k + f^k(d) + \beta^k x_{it})}{1 + \exp(c_i^k + f^k(d) + \beta^k x_{it})}$. In the absence of unobserved heterogeneity ($c_i^k = 0$) one assumes that the exit and re-entry equations represent conditionally independent processes and therefore the log-likelihood function can be maximised separately for all

[42] Lancaster notes “there is nothing to be gained by adopting the more elaborate proportional hazard model over the piecewise, constant one. Essentially, since with grouped data we can know nothing about the way the hazard varies within the interval, the best we can do is to estimate its average level and we might as well work with the simplest model, in which that level is constant”, p.181.

[43] Wooldridge (2002) notes: “usually the duration dummies are unrestricted, in which case x_{it} does not contain an intercept” (p.709). Alternatively, as explained by Jenkins (2001), one can drop one duration dummy to use it as a reference and fit an overall intercept term of the model. In our 12-year data set, the exclusion of left-censored spells implies that exit rates can only be calculated up to a maximum duration of 10 years (se also figures 2 and 3). So D=10 if we do not use an intercept and D=9 if we do use an intercept term.

[44] Appendix II contains a sensitivity analysis on the distribution of the random error term for the estimation results.

spells of a given type k as: $\log L^k = \sum_{i=1}^{N} \sum_{t=1}^{T} m_{it} \{(1-l_{it})\log(1-P_{it}^k) + l_{it}\log(P_{it}^k)\}$ where l_{it} indicates whether an exit from the spell is observed for individual i in t ($l_{it} = 1$) or not ($l_{it} = 0$) and $m_{it} = 1$ if for individual i in t a spell of type k is being observed and $m_{it} = 0$ otherwise.

If the transition equations depend on a random effect that is allowed to be correlated across spells of different types ($c_i^k \neq 0$), poverty and non-poverty spells cannot be treated separately and a simultaneous estimation is necessary. One might expect, indeed, that there are individual-specific unobserved effects like ability, motivation or general attitudes that affect each type of transition. If individuals have a high propensity to leave poverty, one may expect they have also have a low propensity to re-enter poverty. If that would be so, there would be negative correlation among the unobserved effects of the transition equations. In a model that allows for unobserved heterogeneity, an additional problem arises. The probability that the first non-left censored spell is a poverty spell will depend on individual characteristics including unobserved effects, creating an initial conditions problem whereby the identity of the first complete spell we observe is endogenous. To control for the selection bias that may arise I follow Heckman's (1981) approximation method and define a probability of being in a spell of type k at the initial year of observation as a function of individual characteristics and unobserved effects and estimate this probability together with the transition equations while allowing the unobserved effect of the initial condition equation to be correlated with the unobserved effects of the transition equations. To identify the model I use explanatory variables in the initial condition equation that are excluded from the transition equations:

$P_i^0 = \dfrac{\exp(q + \gamma' W_{i0})}{1 + \exp(q + \gamma' W_{i0})}$ where t=0 is used to denote the calendar year in which the first non-left censored poverty spell started, W_{i0} $(\neq x_{it})$ is a vector of observable characteristics and q corresponds to the unobserved effect. If we denote the joint trivariate distribution of the random unobserved effects by $F(c^p, c^{np}, q)$, the log-likelihood function for the whole sample becomes:

$$\log L = \sum_{i=1}^{N} \log\left\{ \int_{R(c^{np})} \int_{R(c^{p})} \int_{R(q)} L_i(c^p, c^{np}, q) dF(c^p, c^{np}, q) \right\}$$

where individual i's contribution equals

$$L_i(c^p, c^{np}, q) = P_{i0}(q)^{np_{i0}} (1 - P_{i0}(q))^{p_{i0}} * \prod_{t=1}^{T} \left\{ (1 - P_{it}^{np}(c^{np}))^{1-l_{it}} (P_{it}^{np}(c^{np}))^{l_{it}} \right\}^{np_{it}} \left\{ (1 - P_{it}^{p}(c^{p}))^{1-l_{it}} (P_{it}^{p}(c^{p}))^{l_{it}} \right\}^{p_{it}}$$

where $p_{it} = 1$ when the spell in t is a poverty spell (and $p_{it} = 0$ otherwise) and if it is a non-poverty spell, $np_{it} = 1$ (and $np_{it} = 0$ otherwise). P_{0i}^{p} denotes the probability of observing a poverty spell in the first year of the observation. In order that the sample likelihood can be estimated[45] a trivariate normal distribution with mean 0 is assumed for the unobserved effects.

One can test the ignorability of the initial condition and of individual-specific time, constant effects on the basis of the correlations of the cross-equations error terms. We write the covariances between the unobserved effects as:

$$\rho_1 = \mathrm{cov}(c_i^p, c_i^{np})$$

$$\rho_2 = \mathrm{cov}(c_i^{np}, q)$$

$$\rho_3 = \mathrm{cov}(c_i^p, q)$$

Where ρ_1 summarizes the association between unobserved effects determining poverty exit and poverty re-entry, ρ_2 the association between unobserved effects determining poverty re-entry and initial poverty status and ρ_3 the association between unobserved effects determining poverty exit and initial poverty status. If these associations are significant, there is evidence of unobserved heterogeneity and of an initial condition problem.

[45] The estimation requires nonlinear optimization methods. The Newton and Raphson technique used (within the PROC GLIMMIX procedure of SAS 9.2.) is a numerical algorithm to find the first-order and second order derivatives of a log-likelihood function. Since with a logistic regression the log-likelihood is globally concave, the function can have at most one maximum (Amemiya, 1985) and there are no problems of local maxima (Allison, 2008).

4.3. Results

The first column of table 6 reports the estimates for separately estimated poverty transitions that do not control for unobserved heterogeneity nor for an initial condition problem. The second column reports the estimates from the joint estimation of exit and re-entry rates allowing for correlated unobserved heterogeneity but not for the initial condition. Including the latter as well leads to the estimates in the third column.

When one wants to control for unobserved heterogeneity and explanatory variables which may be time-varying, it is difficult to relax the requirement they should be strictly exogenous. Since we only are only interested in the coefficients of duration dependence, while controlling for all kind of observed and unobserved effects, we could not include all variables at our disposal and only included explanatory variables that can be justified as strictly exogenous such as age, gender, year[46]. We also consider education level (measured as primary, secondary or high school education) as exogenous for individuals that at the moment they are sampled approach retirement. Being member of the second pillar is considered exogenous since in Belgium this decision is taken at latest at the age of 25 or at the start of the employment relationship and it is taken by the employer and external to the employed worker[47]. Finally, there are time-varying variables that may be exogenous or endogenous like employment status or household composition. We tested for their exogeneity by regressing employment status and household composition on lagged poverty status. Since this was significant, we concluded that these are endogenous and excluded them[48]. The right approach would be to estimate the two poverty transitions

[46]AGE and YEAR are time-varying but can be treated essentially in the same way as time-invariant variables, as explained by Lancaster (1990, p.21).

[47] Since the law of 6th April 1995, the decision to introduce an occupational pension scheme is the exclusive authority of the employer. In addition, the age at which the employee becomes member of a scheme is at maximum 25 years old or at the start of the employment relationship.

[48] In contrast to VanKerm (2004), Makovec (2005), Capellari and Jenkins (2002), Capellari and Jenkins (2004), Capellari (1999), Arulumpalam, Booth and Taylor (1998), Nicaise and Deblander (2005), Hansen and Wahlberg (2004), Devicienti and Gualteri (2007), Andrèn (2007), Poggi (2007) who include in a model with unobserved effects employment status and/or household composition and thus assume the latter are strictly exogenous.

equations (and the initial condition equation) together with an equation for employment status and marital status[49] but this is an issue for future research.

A practical difficulty concerns the choice of exclusion restrictions in the third column. The theoretical idea is clear: good exclusion restrictions should only affect the probability that the first spell of an individual is a poverty spell while it has no effect on poverty transitions. The empirical literature does, however, not provide much guidance on this matter. If a variable is used in the initial condition equation while it is excluded by the researcher from the transition equation although it would be significant in the transition equation, it would lead to measurement error. Therefore, we tested explicitly whether the variables that were introduced in the initial condition equation were insignificant in the transition equation. We dispose of the variable for subjective reported health status that, if it could argued to be time-invariant, could be included as strictly exogenous instrument since it appeared to be strongly significant in the initial condition equation but not in the transition equations. The common practice of all models on poverty dynamics that account for unobserved effects and dispose of a health variable is to assume that this is strictly exogenous: Meghir and Whitehouse (1997), Arulumpalam, Booth and Taylor (2000), Jenkins (2002), Capellari and Jenkins (2004) and Nicaise and Deblander (2005). Nonetheless, it is true that in the cited models health status is measured at the beginning of the sampling period while in our case it is measured at the end of the sampled period (in 2001). If self-reported health status would not be time-constant, it could be not strictly exogenous and lead to inconsistent estimates. For our dataset, there is no way to test whether it is exogenous or endogenous. Survey information[50] seems to indicate, however, that subjective reported health status is rather time-invariant. To avoid discussion, we proxied self-reported health status by using life expectancy measured in the first year of the sampling period 1991 by age, education level and gender[51]. This is strictly exogenous

[49] Biewen (2004) and Aasve et al. (2006) are up to now the only ones that estimate simultaneously employment status and household composition in a model of poverty dynamics.

[50] Kington and Smith (1998): "Self-reported health status is not used to measure temporary health problems but to include general physical, social and emotional function. Health in old age reflects one's long-term health history. The study's findings show that health status in advanced years is greatly influenced by a history of health that goes back to one's childhood and reaches even beyond personal health status to include the health status of parents and siblings throughout their lives."

[51] Deboosere and Gadeyne (2002).

and has the advantage of containing more variability than the usual instrument; unemployment rate. Finally, after also experimenting with growth rate of GDP, a year dummy for 1992 and combinations of instruments, the best fit and most significant results were obtained with unemployment rate and health expectancy of the head as instruments.

Table 6: multiple-spell discrete-time hazard model with unobserved effects[52]			
Variable	Without heterogeneity Without initial condition	With heterogeneity Without initial condition	With heterogeneity with initial condition
Exit			
Intercept	- 17.426** (1.35)	- 16.622** (1.40)	- 16.307** (1.51)
Duration 1 year	6.278** (0.26)	6.321** (0.27)	5.218** (0.28)
Duration 2 years	5.351** (0.28)	5.451** (0.28)	4.519** (0.28)
Duration 3 years	5.084** (0.28)	5.210** (0.28)	4.378** (0.28)
Duration 4 years	5.117** (0.28)	5.248** (0.28)	4.491** (0.28)
Duration 5 years	4.909** (0.29)	5.055** (0.29)	4.373** (0.29)
Duration 6 years	4.727** (0.29)	4.879** (0.30)	4.257** (0.30)
Duration 7 years	4.535** (0.31)	4.697** (0.31)	4.128** (0.31)
Duration 8 years	4.395** (0.32)	4.557** (0.32)	4.024** (0.33)
Duration 9 years	4.409** (0.32)	4.659** (0.34)	4.163** (0.35)
AGE	0.399** (0.04)	0.374** (0.04)	0.400** (0.05)
AGE squared	- 0.003** (0.0003)	- 0.003** (0.0003)	- 0.003** (0.0004)
Female head	- 0.425** (0.04)	- 0.436** (0.04)	- 0.487** (0.05)
Member second pillar	0.350** (0.03)	0.376** (0.04)	0.429** (0.04)
Education level head			
Low secondary general	- 0.052 (0.06)	- 0.053 (0.06)	- 0.075 (0.07)
Low secondary technical	0.184* (0.06)	0.213** (0.06)	0.230** (0.07)
Low secondary professional	0.0727 (0.06)	0.077 (0.06)	0.069 (0.07)
High secondary general	0.116	0.135	0.144

[52] Since the sample consists of repeated observations on the same household, standard errors are adjusted to account for the dependence at the level of the household.

	(0.07)	(0.08)	(0.09)
High secondary technical	0.207** (0.07)	0.218 (0.08)	0.236 (0.09)
High secondary professional	- 0.112 (0.09)	- 0.124 (0.10)	- 0.160 (0.11)
University	0.298** (0.05)	0.310** (0.05)	0.341** (0.06)
Education level spouse of household head			
Low secondary general	- 0.090 (0.06)	- 0.124 (0.06)	- 0.141 (0.07)
Low secondary technical	- 0.052 (0.103)	- 0.069 (0.11)	- 0.068 (0.12)
Low secondary professional	0.026 (0.06)	0.0165 (0.06)	0.028 (0.07)
High secondary general	0.056 (0.08)	0.065 (0.09)	0.086 (0.10)
High secondary technical	0.073 (0.13)	0.059 (0.14)	0.079 (0.16)
High secondary professional	- 0.028 (0.09)	- 0.039 (0.10)	- 0.034 (0.12)
university	0.219** (0.085)	0.227** (0.09)	0.278** (0.10)
Re-entry			
intercept	- 8.882** (1.66)	- 8.209** (1.77)	- 7.659** (1.88)
Duration 1 year	5.882** (0.33)	5.935** (0.33)	5.450** (0.33)
Duration 2 years	5.183** (0.33)	5.347** (0.33)	4.990** (0.33)
Duration 3 years	4.868** (0.34)	5.115** (0.34)	4.842** (0.34)
Duration 4 years	4.564** (0.34)	4.873** (0.34)	4.665** (0.34)
Duration 5 years	4.303** (0.34)	4.672** (0.35)	4.522** (0.35)
Duration 6 years	4.213** (0.35)	4.630** (0.35)	4.531** (0.36)
Duration 7 years	4.175** (0.36)	4.634** (0.36)	4.581** (0.37)
Duration 8 years	3.856** (0.38)	4.355** (0.39)	4.348** (0.39)
Duration 9 years	3.945** (0.41)	4.453** (0.41)	4.476** (0.42)
AGE	0.088 (0.05)	0.065 (0.05)	0.061 (0.05)
AGE SQUARED	- 0.0009** (0.0004)	- 0.0007** (0.0005)	- 0.0007 (0.0005)
Female head	0.013 (0.05)	0.039 (0.06)	0.071 (0.07)
Member second pillar	- 0.409** (0.05)	- 0.478** (0.05)	- 0.527** (0.05)
Education household head			
Low secondary general	0.065 (0.08)	0.101 (0.08)	0.124* (0.09)
Low secondary technical	- 0.129	- 0.167	- 0.189

	(0.07)	(0.08)	(0.09)
Low secondary professional	- 0.0058 (0.07)	- 0.026 (0.08)	- 0.033 (0.08)
High secondary general	- 0.101 (0.08)	- 0.142 (0.09)	- 0.156 (0.10)
High secondary technical	- 0.055 (0.09)	- 0.106 (0.10)	- 0.137 (0.11)
High secondary professional	0.168 (0.10)	0.183 (0.11)	0.197 (0.12)
university	- 0.313** (0.07)	- 0.351** (0.07)	- 0.381** (0.08)
Education level spouse			
Low secondary general	0.221** (0.07)	0.280** (0.08)	0.316** (0.09)
Low secondary technical	- 0.045 (0.12)	- 0.041 (0.041)	- 0.033 (0.15)
Low secondary professional	- 0.024 (0.08)	- 0.006 (0.08)	- 0.002 (0.09)
High secondary general	0.046 (0.11)	0.080 (0.123)	0.102 (0.13)
High secondary technical	- 0.025 (0.15)	0.005 (0.16)	0.0039 (0.18)
High secondary professional	0.142 (0.11)	0.192 (0.12)	0.210 (0.14)
university	- 0.358 (0.10)	- 0.366 (0.11)	- 0.388 (0.12)
Initial condition[53]			
Life expectancy head			- 5.199** (1.33)
Unemployment rate			21.605** (7.25)
Covariance parameter estimates of unobserved effects			
		ρ_1 =, 0.59 (0.04)	ρ_1 =, 0.95 (0.05) ρ_2 =0.105 (0.04) ρ_3 =, 0.035 (0.04)
Formal likelihood ratio tests of significance of covariance parameter estimates			
		H_0: ρ_1 =0 χ_1^2 =9987;Pr> χ_1^2 <0.0001	H_0: ρ_1 =0 χ_1^2 =293.89;Pr> χ_1^2 <0.0001 H_0: ρ_2 =0 χ_1^2 =3.45;Pr> χ_1^2 =0.05 H_0: ρ_3 =0 χ_1^2 =0.40;Pr> χ_1^2 =0.52 H_0: $\rho_1 = \rho_2$ =0 χ_2^2 =366.33;Pr> χ_1^2 <0.0001

[53] The initial condition equation also includes all the explanatory variables in the poverty transition equations but these are not displayed to save space.

			H$_0$: $\rho_2 = \rho_3 = 0$ χ^2_2 =3.45;Pr> χ^2_1 =0.17 H$_0$: $\rho_1 = \rho_2 = \rho_3 = 0$ χ^2_3 =400.25;Pr> χ^2_1 <0.0001
- 2LogLikelihood	461013	459184	405304
Number observations	72844	72844	72844
Residual	1.06 (0.005)	0.97 (0.006)	0.86 (0.005)
* denotes significance at 5% level, ** denotes significance at 1% level; standard errors in parenthesis. The reference person is a male head with primary school education, no member of second pillar, living in Antwerp, with spouse with primary school education in 1992. Year and province dummies are included in all equations.			

We primarily compare the coefficients of duration dummies across the three scenarios. When taking into account unobserved heterogeneity, the coefficients become larger. However the duration dummies remain strongly significant and decrease with duration suggesting a genuine causal effect of duration dependence. The unobserved effects do not really change the coefficients of the individual characteristics which means that the unobserved effects are not correlated with the already included individual characteristics. As is typically the case, taking into account the sample selection bias, when going from the second to the third column, there is a reduction in the estimated coefficients: the exogeneity hypothesis leads to over-estimate both size and significance of the estimated coefficients.

Most of the individual characteristics that are significant, like education level, female head and membership of second pillar, have opposite signs in the exit and re-entry equations. As in Stewart and Swaffield (1999), Capellari and Jenkins (2002) and VanKerm (2004) the unobserved effects that lead to poverty exit are also decreasing poverty re-entry since ρ_1 has a negative sign (ρ_1=, 0.59) and is strongly significant. This means that besides the observable characteristics that reduce exit and increase re-entry there are also, in addition, unobserved effects that lead to the same kind of persistence in poverty. This remains so if we also take the initial condition into account (ρ_1=, 0.95).

The positive sign of ρ_2 indicates that the unobserved effects that mean that individuals are likely to be initially poor are also increasing the risk of poverty exit. This sign is interpreted by Stewart and Swaffield (1999) and Jenkins and Capellari (2002) as follows:

given that it measures the correlation between the probability of having a poverty transition and being initially poor, the negative sign is analogous to a negative coefficient in the regression of poverty transitions on poverty status, i.e. Galtonian regression towards the mean. Finally, ρ_3 has a negative sign meaning that the unobserved effect that means that individuals are likely to be initially poor is also decreasing the risk of poverty re-entry but is not significant.

A formal likelihood ratio test[54] of significance of the covariance parameter estimates confirm that the model that allows for correlation between the unobserved effects is clearly to be preferred to the one that does not; what we interpret as evidence of unobserved heterogeneity and an initial condition problem. It also shows that the hypothesis that only ρ_3=0 cannot be rejected.

Up to now, we assumed the population is a group of homogenous individuals. Appendix I and III repeat this analysis but separately for the elderly below 65 and above 65 years old and separately for the elderly covered under the social security system of the employed and self-employed respectively.

5. Concluding remarks

The matching of the National Register with the Income Tax Returns and Socio-economic Survey provided evidence of strong income mobility: 1) Every year about 14% of the Belgian civil population is out of the Income Tax Returns, while only 4.9% of households do not appear for any year 1991-2002 in the Income Tax Returns; 2) 37% of the Belgian elderly experience poverty once over a period of 12 years, which is much larger than the 12% of Belgian elderly who are poor in a given year.

[54] The likelihood ratio statistic is formed as twice the difference of the log likelihoods of the unrestricted model and the restricted model which has a χ^2_Q distribution under the null hypothesis H_0 with Q the number of restrictions imposed.

About 30% of those who become poor leave poverty after one year and are only *transitory* poor. The bulk of the elderly poor are, however, *persistently* poor. The question arises whether this persistence in poverty is true or spurious. The estimation of a multiple spell discrete-time hazard model, controlling for unobserved effects and a significant initial condition problem, showed a genuine causal effect of duration dependence. One does not know a lot of the mechanism that lies behind genuine persistence in poverty. It has been suggested that persistence may be due to depreciation of human capital or adverse work incentives. The latter illustrates the poverty trap: people may be given a financial incentive not to work while at the same time they slip into poverty. This suggestion sounds reasonable since in Belgium elderly unemployed are exempted from the search for a job and thus easily exposed to depreciation of human capital and employers are reluctant to invest in the human capital of elderly workers. In addition, in Belgium both employers and the government design retirement pathways that give elderly strong incentives to leave the labour market as soon as possible.

Bibliography

Aasve A., Burgess S., Dickson M., Propper C. (2006), "Modelling poverty by not modelling poverty: an application of a simultaneous hazards approach to the UK", Working Paper 2005, 26, ISER, University of Essex.

Allison P.D. (1984), *Event history analysis. Regression for longitudinal event data,* Sage University Paper Series (Quantitative applications in social sciences).

Allison P.D. (2008), "Convergence failures in logistic regressions", SAS Global Forum 2008, University of Philadelphia.

Amemiya T. (1985), *Advanced econometrics*, Harvard University Press, Cambridge.

Andren T. (2007), "The persistence of welfare participation", Discussion Paper 3100, IZA, Bonn.

Antolin P., Dang T., T., Oxley H. (1999), "Poverty dynamics in four OECD countries", Economics Department Working Paper 212, OECD, Paris.

Arranz J.M., Cantto O. (2006), "The dynamics of poverty in Spain: the relevance of considering multiple, spells and their accumulation", Discussion Paper, University of Alcala.

Arulampalam W., Booth A., Taylor M. (2000), "Unemployment persistence", *Oxford Economic papers*, Oxford University Press, 52 (1), 24, 50.

Bane M.J, Ellwood D.T. (1986), "Slipping into and out of poverty: the dynamics of spells", *Journal of Human Resources* 21 (1), 1, 23.

Bardasi E., Jenkins S., Rigg J. (2000), "Retirement and the economic well, being of the elderly: a British perspective", ISER, University of Essex.

Biewen M. (2003), "Who are the chronic poor? Evidence on the extent and the composition of chronic poverty in Germany", Discussion Paper 350, DIW Berlin.

Biewen M. (2004), "Measuring state dependence in individual poverty status: are there feedback effects to employment decisions and household composition?", Discussion paper 1138, IZA Bonn.

Blöndal S., Scarpetta S. (1998), "The retirement decision in OECD countries", Economics Department Working Paper 202, OECD, Paris.

Boldrin M., Dolado J. J. Jimeno J.F., Perrachi F., (1999), "The Future of Pension in Europe", *Economic Policy*, 29, 289, 320.

Bourreau, Dubois C., Jeandidier B., Berger F. (2003), "Poverty dynamics, family events, labour market events in Europe: are there any differences between women and men?, The 2003 Conference of the European Panel Users Network, University of Essex.

Burniaux J., M., Padrini F., Brandt N. (2007), "Labour market performance, income inequality and poverty in OECD countries", Economics Department Working Paper 500, OECD, Paris.

Callens M., Croux C., Avramov D. (2003), "Poverty dynamics in Europe: a multi, level discrete, time recurrent hazard analysis", EPUNET 2005 Conference, University of Essex.

Cantillon B., De Lathouwer L., Marx I., Van Dam R., Van den Bosch K. (1999), "Sociale indicatoren 1976, 1997", CSB bericht, University of Antwerp.

Capellari L., Jenkins S. (2002), “Modelling low income transitions”, Working Paper 2002, 8, ISER, University of Essex.

Capellari L., Jenkins S. (2004), “Modelling low pay transition probabilities, accounting for panel attrition, non, response and initial conditions”, Working Paper 2004, 08, ISER, University of Essex.

Cappellari L., Dorsett R., Haile G. (2007), “State dependence, duration dependence and unobserved heterogeneity in the employment transition over the 50’s”, Working Paper 2007, 16, ISER, University of Essex.

Casey B., Yamada A. (2002), “Getting older, getting poorer? A study of the earnings, pensions, assets and living arrangements of older people in nine countries”, Occasional Papers 60, OECD, Paris.

Cockx B. (1998), “Analysis of transition data by the minimum chi, square method : an application to welfare spells in Belgium”, *The Review of Economics and Statistics* 79 (3), 392, 405.

Deboosere P., Gadeyne S. (2002),”Sterftetafels naar geslacht, gewest en onderwijsniveau in België, 1991, 1996”, Working Paper Steunpunt Demografie , Vakgroep Sociaal onderzoek, Free University of Brussels.

Dejemeppe M., Cockx B. (2005), “Duration dependence in the exit rate out of unemployment. Is it true or spurious”, *Journal of Applied Econometrics* 20 (1), 1, 23.

Deleeck H., Cantillon B., Meulemans B., Van den Bosch K. (1992), “Some longitudinal results of the belgian socio, economic panel”, *Journal of Income Distribution* 2 (2), 211, 231.

Dejemeppe M., Cockx B. (1998), “La conception des politiques en faveur de l’emploi. L’importance d’un diagnostic des causes du chômage structurel”, Treizième Congrès des économistes de langue française, CIFoP, Charleroi.

Dewilde C. (2004), “Poverty mobility in the belgian and british welfare regimes : the impact of demographic and labour market events”, Research Group on Poverty, Social Exclusion and the City, University of Antwerp.

Devicienti F. (2002), “Estimating poverty persistence in Britain”, Working Paper 1, Laboratorio R.Revelli, Centre for Employment Studies, Moncalieri.

Devicienti F. (2001a), “Poverty persistence in Britain: a multivariate analysis using the BHPS, 1991, 1997”, *Journal of Economics* 9, 1, 34.

Devicienti F., Gualtieri V. (2007), “The dynamics and persistence of poverty: evidence from Italy”, Working Paper 63, Laboratorio R. Revelli, Centre for Employment Studies, Moncalieri.

Dorn D., Souza A. (2005), “Early retirement: free choice or forced decision?”, Working paper 1542, CESifo.

Fenge R., Pestieau P. (2005), *Social security and early retirement*, MIT Press, Cambridge.

Fertig M., Tamm M. (2007), ”Does money buy higher scholling? Evidence from secondary school track choice in German”, Working Paper 2007, 58, ECINEQ.

Finnie R., Sweetman A. (2002), “Poverty dynamics: empirical evidence for Canada”, School of policy studies, Working paper 29, Queens University.

Fouarge D., Layte R. (2003), “Duration of poverty spells in Europe”, Working Paper 2003, 47, EPAG.

Fouarge D., Muffels R. (2003) “The role of European welfare states in explaining resources deprivation”. Working Paper 2003, 41, EPAG.

Förster M. (2000), “Trends and driving factors in income distribution and poverty in the OECD area”, Labour market and Social Policy Occasional papers 42, OECD, Paris.

Franck (1987), “La fraude fiscale et la sous, estimation fiscale et leur impact sur la distribution des revenues”, in Ginsburgh V., Pestieau P. (eds.), *L’économie informelle : Fraude fiscale, Travail en noir et autres activités non déclarées*, 23, 25, éditions Labor, Brussels.

Grüber J., Wise D. (eds.) (1999), *Social security and retirement around the world* . University of Chicago Press, Chicago.

Grüber J.,Wise D. (eds.) (2004), *Social security programs and retirement around the world: Micro, estimation*, University of Chicago Press, Chicago.

Hansen J., Wahlberg R. (2004), “Poverty persistence in Sweden”, Discussion Paper 1209, IZA, Bonn.

Heckman J. (1981), “The incidental parameters problem and the problem of initial conditions in estimating a discrete, time discrete data stochastic process”, in Manski, C.F., McFadden (eds.), S*tructural analysis of discrete data with economic applications*, MIT Press, Cambridge.

Heckman J., Singer B. (1984), “A method for minimizing the impact of distributional assumptions in econometric models for duration data”, *Econometrica* 52 (2), 271, 320.

Jenkins S. (2000), “Modelling household income dynamics”, *Journal of Population dynamics 13*, 529, 567

Jenkins S., Rigg J. (2001), *The Dynamics of Poverty in Britain*, Department for Work and Pensions Research Report 157, Corporate Document Services, Leeds.

Jenkins S. (1995), “Easy estimation methods for discrete, time duration models”, *Oxford bulletin of economics and statistics*, 57, 1, 129, 138.

Jenkins S. (2005), *Survival analysis*, ISER, University of Essex.

Jousten A. (2005), “The benefits of separating early retirees from the unemployed: simulation results for belgian wage earners”, Discussion Paper 5077, CEPR.

Jousten A., Lefebvre M., Perelman S., Pestieau P. (2008), “The effects of youth unemployment on early retirement”, Working Paper 08/30, IMF.

Kington R., Smith J. (1998), “The relationship between socio, economic status and health of elderly”, Research Brief 5020 Labor and Population, RAND.

Lancaster T. (1990), *The econometric analysis of transition data*, Econometric Society Monographs, Cambridge University Press, Cambridge.

Lillard L.A., Willis R.J. (1978), “Dynamic aspects of earnings mobility”, *Econometrica* 46 (5), 985, 1012.

Lillard L.A., Panis C. (1998), “Panel Attrition from the Panel Study of Income dynamics: Household Income, Marital Status, and Mortality”, *The Journal of Human Resources* 33 (2), 437, 457.

Lindeboom,M. (1998), “Microeconometric analysis of the retirement decision: the Netherlands”, Economics Department Papers 207, OECD, Paris.

Makovec M. (2005), “Poverty entry and exit during old, age: comparative evidence from the ECHP”, European Centre for Social Welfare Policy and Research.

McKernan S.M., Ratcliffe C. (2002), “Transition events in the dynamics of poverty”, The Urban Institute, Washington.

Meghir C., Whitehouse E. (1997), “Labour market transitions into retirement”, *Journal of Econometrics* 79, 327, 354.

Meyer B. (1990), “Unemployment insurance and unemployment spells”, *Econometrica* 58 (4), 757, 782.

National Office for Pensions (2008), “Jaarlijkse statistiek van uitkeringsgerechtigden: toestand op 1 januari 2007”, Brussels.

Nicaise I., DeBlander R. (2005), “Maatschappelijke keuzen, structurele armoede en sociale kost”, HIVA, University of Leuven.

Nicoletti C., Rondinelli C. (2006), “The (mis)specification of discrete, time duration models with unobserved heterogeneity: a Monte Carlo study”, Working Paper 2006, 53, ISER, University of Essex.

Nickell S.J. (1979), “The effect of unemployment and related benefits on the duration of unemployment”, Economic Journal 89 (353), 34, 49.

OECD (1994), *The OECD Jobs Strategy*, OECD, Paris.

OECD (1997), *Implementing the OECD Jobs strategy: Lessons from Member Countries’ experiences*, OECD, Paris.

OECD (2005a), *Extending opportunities: how active social policy can benefit us all*, OECD, Paris.

OECD (2005b), *Increasing financial incentives to work: the role of employment conditional benefits*, Employment Outlook, OECD, Paris.

Otto B., Siedler T. (2003), “Poverty in Western and Eastern Germany – A detailed comparison”, *Economic Bulletin* 2/2003, 71, 76.

Perelman S., Schleiper A., Stevart M. (1998), “Dix années plus tard, d’un congrès à l’autre: l’apport des statistiques fiscales à l’étude de la distribution des revenues”, Treizième Congrès des économistes de langue française, CIFoP, Charleroi.

Poggi A. (2007), “Does persistence of social exclusion exist in Spain?”, *Journal of Economic Inequality 5,* 53, 72.

Stevens A.H. (1999), “Climbing out of poverty, falling back in: measuring the persistence of poverty over multiple spells”, *Journal of Human Resources* 3, 557, 588.

Stewart M. (2007), “The inter, related dynamics of unemployment and low pay”, *Journal of applied econometrics* 22 (3), 511, 531.

Stewart M., Swaffield J.K. (1999), “Low pay dynamics and transition probabilities”, *Economica 6*6, 23, 42.

Sueyoshi G. (1995), “A class of binary response models for grouped duration data”, *Journal of applied econometrics* 10, 411, 431.

Valletta R. (2005), “Les entrées et les sorties de la pauvreté dans les économies avancées : la dynamique de la pauvreté au Canada, en Allemagne, en Grande, Bretagne et aux Etats, Unis", Série de documents de recherche – Division de la statistique du Revenu, Federal Reserve Bank, Canada.

van den Berg G. (2001), "Duration models : specification, identification and multiple durations", in Heckman J., Leamer E. (eds.) *Handbook of Econometrics,* Volume V, North, Holland.

Van Kerm P. (2004),"Une évaluation économétrique des flux vers et hors de la pauvreté en Belgique", IRISS Working Paper Series 2004, 04, CEPS/INSTEAD.

Wooldridge J. (2002), *Econometric analysis of cross section and panel data,* MIT Press, Cambridge.

Zaidi A., Makovec M., Fuchs M., Lipszyc B.,Lelkes O., Rummel M., Marin B., DeVos K. (2006), "Poverty of elderly people in EU, 25", First Report for the Project "Poverty of elderly people", financed by the European Commission.

Zoyem J.P. (2002), "La dynamique des bas revenus : une analyse des entrées, sorties de pauvreté", Document de travail 2002/11, INSEE.

Appendix I: Sensitivity of the results to age groups above and below 65 years old

Usually studies that describe the phenomenon of poverty distinguish between the working age population and the non-working age population where the former groups people below 65 years old and the latter those above 65 years old. The cut-off age, 65, corresponds to the age at which entry to the old-age pension system is almost 100% in most OECD countries. Nonetheless, the average retirement age for men was about 57 and for women 54 in Belgium in 2001. This suggests that the age of 65 corresponds purely to the age at which the elderly unemployed, disabled or early retired are automatically switched from one social security system to another. It is also the age after which labour market attachment is expected to be absent or extremely low[55]. We investigated whether the same statistical models were appropriate for individuals who were less than 65 at the beginning of their first non-left-censored spell and above 65 at the beginning of their first non-left-censored spell. To check this, we estimated exit and re-entry rates simultaneously while controlling for observed and unobserved heterogeneity and without left-censored spells for both age groups separately.

Table 7: Multiple-spell discrete-time hazard model with unobserved effects for different age groups[56]

Variable	Below 65 years old at start first non-left censored spell	Above 64 years old at start first non-left-censored spell
Exit		
Intercept	- 19.100** (1.74)	- 15.768 (62.51)
Duration 1 year	6.198** (0.27)	5.410** (0.48)
Duration 2 years	5.358** (0.28)	4.317** (0.53)
Duration 3 years	5.124** (0.28)	3.746** (0.57)
Duration 4 years	5.175** (0.28)	3.450** (0.63)
Duration 5 years	4.967** (0.29)	3.369** (0.66)
Duration 6 years	4.802**	2.739**

[55] Jenkins and Rigg (2001) also estimate separate models for elderly above 60 and individuals between 0-60 years old arguing that "the association between labour market attachment and hazard rates may differ for elderly and non-elderly households".

[56] Since the sample consists of repeated observations on the same household, standard errors are adjusted to account for the dependence at the level of the household.

	(0.30)	(0.84)
Duration 7 years	4.558** (0.31)	3.796** (0.77)
Duration 8 years	4.442** (0.33)	/
Duration 9 years	4.524** (0.35)	/
AGE	0.463** (0.35)	0.296 (1.78)
AGE squared	- 0.004** (0.0005)	- 0.002 (0.01)
Female head	- 0.444** (0.05)	- 0.392 (0.25)
Member second pillar	0.403** (0.04)	0.314* (0.15)
Education level head		
Low secondary general	0.013 (0.06)	- 0.756* (0.36)
Low secondary technical	0.264** (0.06)	- 0.803 (0.46)
Low secondary professional	0.152* (0.06)	- 0.754* (0.33)
High secondary general	0.231** (0.08)	- 0.900* (0.39)
High secondary technical	0.277** (0.08)	- 0.442 (0.65)
High secondary professional	- 0.052 (0.10)	- 0.800 (0.706)
University	0.378** (0.06)	- 0.591* (0.267)
Education level spouse		
Low secondary general	- 0.136* (0.07)	0.168 (0.27)
Low secondary technical	- 0.060 (0.11)	0.233 (0.48)
Low secondary professional	0.037 (0.07)	0.045 (0.28)
High secondary general	0.078 (0.09)	0.197 (0.57)
High secondary technical	- 0.010 (0.14)	1.805 (0.50)
High secondary professional	- 0.031 (0.10)	- 0.08 (0.42)
University	0.205* (0.09)	0.990** (0.36)
Re-entry		
Intercept	- 11.601** (2.11)	- 183,10* (87,39)
Duration 1 year	5.830** (0.33)	5.132** (0.75)
Duration 2 years	5.240** (0.33)	4.545** (0.78)
Duration 3 years	4.994** (0.34)	4.059** (0.83)
Duration 4 years	4.735**	4.093**

	(0.34)	(0.87)
Duration 5 years	4.524** (0.35)	3.890** (0.93)
Duration 6 years	4.463** (0.35)	4.079** (0.98)
Duration 7 years	4.507** (0.36)	2.859** (1.36)
Duration 8 years	4.160** (0.39)	/
Duration 9 years	4.322** (0.42)	/
AGE	0.195** (0.07)	4.769* (2.49)
AGE SQUARED	- 0.001** (0.0006)	- 0.032 (0.01)
Female head	0.016 (0.06)	0.318 (0.28)
Member second pillar	- 0.445** (0.05)	- 0.652** (0.22)
Education level head		
Low secondary general	0.038 (0.08)	1.268** (0.405)
Low secondary technical	- 0.228** (0.08)	0.957** (0.42)
Low secondary professional	- 0.068 (0.08)	0.716 (0.41)
High secondary general	- 0.218* (0.10)	1.057** (0.44)
High secondary technical	- 0.128 (0.10)	1.411 (1.13)
High secondary professional	0.138 (0.11)	1.556* (0.88)
University	- 0.421** (0.07)	0.710* (0.36)
Education level spouse		
Low secondary general	0.267** (0.08)	0.374 (0.40)
Low secondary technical	- 0.030 (0.14)	- 0.810 (0.63)
Low secondary professional	- 0.004 (0.08)	- 0.214 (0.43)
High secondary general	0.082 (0.12)	0.483 (0.54)
High secondary technical	- 0.003 (0.17)	- 0.688 (0.75)
High secondary professional	0.157 (0.13)	0.473 (0.55)
University	- 0.400** (0.11)	- 0.551 (0.51)
Covariance parameter estimates	ρ_1 =- 0.466 (0.04) Residual:0.995 (0.006)	ρ_1 =- 2.434 (0.28) Residual: 0.728 (0.014)

Formal likelihood ratio test of significance of covariance parameter estimates: H_0: ρ_1=0		
	χ_1^2 =8269 (pr> χ_1^2 :<0.0001)	χ_1^2 =1063 (pr> χ_1^2 :<0.0001)
- 2logLikelihood	415088	43065
Number observations	66373	6471
* denotes significance at 5% level, ** denotes significance at 1% level; standard errors in parenthesis. The reference person is a male head with primary school education, is no member of second pillar, with a spouse with primary school education, living in province of Antwerp in 1992. Year and province dummies were included in all equations.		

For the age group below 65 years, all the determinants like university degree or technical school degree, membership of second pillar are highly significant in explaining poverty exit and at the same time decrease poverty re-entry. Being a female head on the other hand decreases exit and increases re-entry, and these are thus likely to be persistently poor. It can be checked that the probability of exit decreases from age 58 on while the probability of re-entry increases with age: with age the probability of being persistently poor increases. Note that for the age group above 65 years old, the number of observations is rather low. This is in the first place because 70% of our sample is between 50-65 years old. In the second place, this is because we deleted left-censored spells and with age elderly experience less poverty transitions. The observed and predicted exit and re-entry rates for both age groups are plotted below.

Figure 5: Observed and predicted exit rates for age groups above and below 65

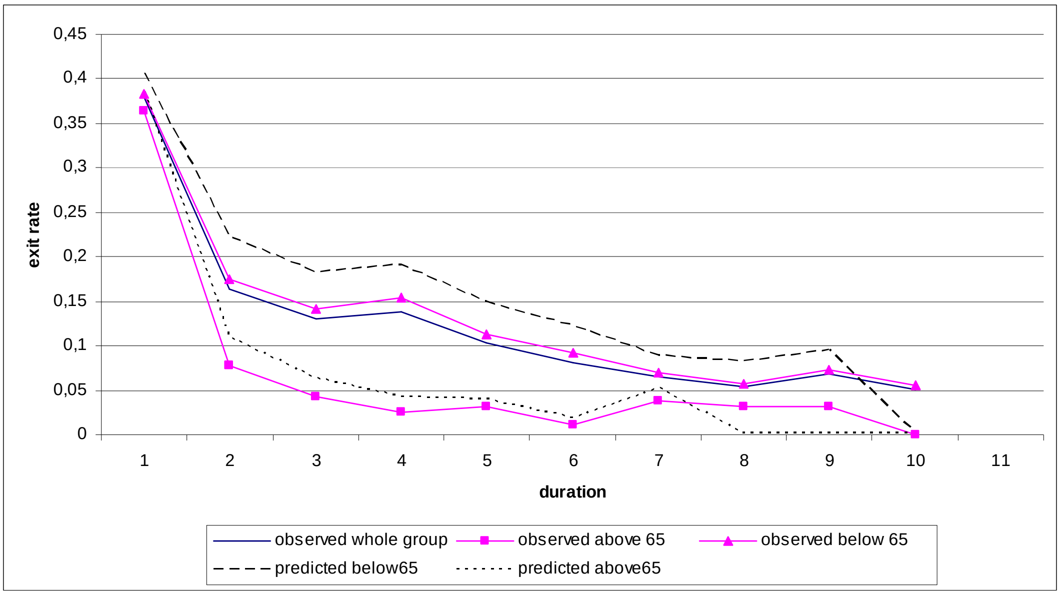
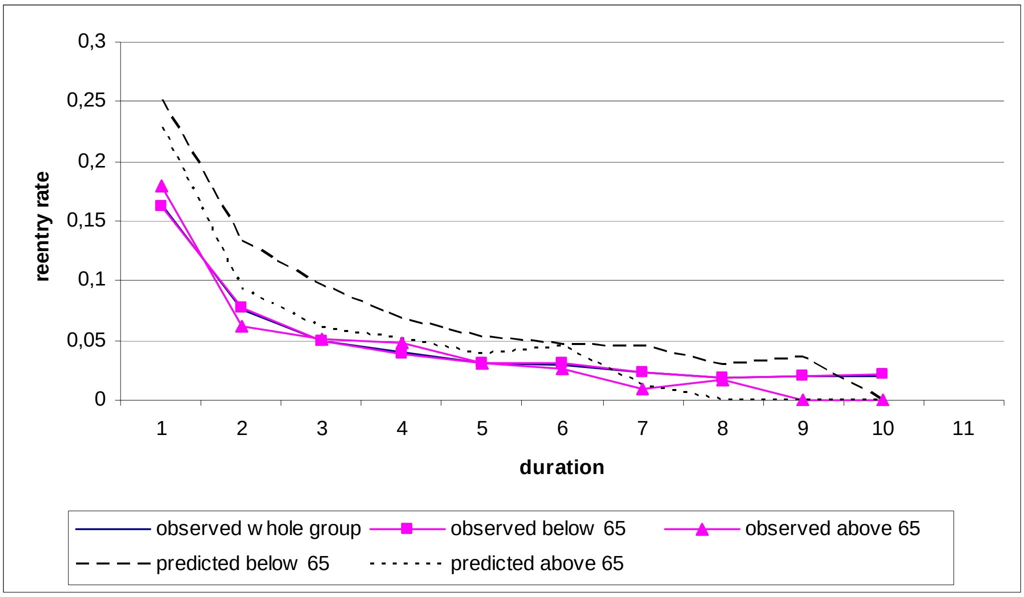


Figure 6: Observed and predicted re-entry rates for age groups above and below 65



As is usual in these models, for each age group the effect of taking into account observed and unobserved heterogeneity leads to an upward shift of hazard rates. One also notes that the observed re-entry rates are almost the same across age groups while, when taking into account heterogeneity, predicted re-entry rates seem to diverge. This suggests that one could correct for more heterogeneity among the age group below 65 than above. The observed exit rates are much higher for the age group below 65 than the age group above 65 and again when taking into account heterogeneity the hazard rates diverge even more. We see that of those whose poverty spell starts after the age of 65[57] there is almost no chance of exiting poverty. This might be explained by the fact[58] that the implicit tax on continued activity after the cut-off age of 65 is, in Belgium, considerably higher than before 65 and one may thus expect that the labour attachment after 65 and the chance to return to the labour market is much lower after 65. This is consistent, in an extreme version, with the idea that high work disincentives are associated with an increased risk of poverty persistence.

---

[57] This may happen, for example if old-age pensions are only indexed to prices while the real poverty line follows wage increase of the whole economy or if a spouse becomes a widow.

[58] Gruber and Wise (1999), Dellis, Jousten and Perelman (1999).

Appendix II: Sensitivity of results to distribution of random error term: extreme value instead of logistic distribution

As noted above, there exist studies on poverty dynamics that use a logistic, normal or extreme value distribution for the random error term. Stevens (1999) and Devicienti (2002) use a logistic distribution for the random error term. Meghir and Whitehouse (1997), Fertig and Tamm (2007), Biewen (2003) and Hansen and Wahlberg (2004) use a normal distribution. Bardasi, Jenkins and Rigg (2001) use an extreme value distribution. Sueyoshi (1995) explores the implications of these specifications for hazard behaviour and notes that “practical experience with discrete, choice models suggests that the predicted probabilities and hence the goodness-of-fit tests for the models will generally be quite similar.” Apart from goodness-of-fit tests, “results from the logit and proportional hazard specifications will be quite similar. In contrast, estimates from a probit-type group duration model should depart significantly from both of these specifications, exhibiting covariate effects that are decidedly non-proportional” while “logistic models are only slightly less proportional than the extreme value specification”. The fact that extreme value and logistic estimation give very similar results is because, if h denotes the hazard rate, the odds ratios (the exponentiated coefficients of the logistic model) will tend, if the hazard is sufficiently small, to the hazard ratio where the latter corresponds to the exponentiated coefficients of the extreme value model. To check whether the coefficients of the extreme value and logistic model differ a lot, we estimated again exit and re-entry rates simultaneously while controlling for observed and unobserved heterogeneity and without left-censored spells in table 8.[59]

[59] Unfortunately, the simultaneous estimation with initial condition did not converge under an extreme value distribution for the random error term.

Table 8: Multiple-spell discrete-time hazard model with unobserved effects for logistic and extreme value distribution of random error term[60]		
Variable	Logistic	Extreme value
Exit		
Intercept	- 16.622** (1.40)	- 14.431** (1.16)
Duration 1 year	6.321** (0.27)	6.005** (0.27)
Duration 2 years	5.451** (0.28)	5.264** (0.28)
Duration 3 years	5.210** (0.28)	5.038** (0.28)
Duration 4 years	5.248** (0.28)	5.064** (0.28)
Duration 5 years	5.055** (0.29)	4.867** (0.28)
Duration 6 years	4.879** (0.30)	4.691** (0.29)
Duration 7 years	4.697** (0.31)	4.497** (0.30)
Duration 8 years	4.557** (0.32)	4.363** (0.32)
Duration 9 years	4.659** (0.33)	4.483** (0.33)
AGE	0.374** (0.04)	0.294** (0.03)
AGE squared	- 0.003** (0.0003)	- 0.0027** (0.0003)
Female head	- 0.436** (0.05)	- 0.360** (0.04)
Member second pillar	0.376** (0.04)	0.311** (0.03)
Education level head		
Low secondary general	- 0.053 (0.05)	- 0.053 (0.05)
Low secondary technical	0.213** (0.06)	0.168** (0.05)
Low secondary professional	0.077 (0.06)	0.060 (0.05)
High secondary general	0.135 (0.08)	0.107 (0.06)
High secondary technical	0.218** (0.08)	0.170** (0.06)
High secondary professional	- 0.124 (0.10)	- 0.124 (0.08)
University	0.310** (0.05)	0.234** (0.04)
Education level spouse		
Low secondary general	- 0.124	- 0.102

[60] Since the sample consists of repeated observations on the same household, standard errors are adjusted to account for the dependence at the level of the household.

	(0.05)	(0.05)
Low secondary technical	- 0.060 (0.08)	- 0.060 (0.08)
Low secondary professional	0.016 (0.06)	0.012 (0.06)
High secondary general	0.065 (0.095)	0.064 (0.075)
High secondary technical	0.059 (0.14)	0.053 (0.11)
High secondary professional	- 0.039 (0.10)	- 0.030 (0.09)
University	0.227** (0.09)	0.188** (0.07)
	Re-entry	
Intercept	- 8.209** (1.77)	- 8.853** (1.57)
Duration 1 year	5.930** (0.33)	5.780** (0.35)
Duration 2 years	5.347** (0.33)	5.242** (0.33)
Duration 3 years	5.115** (0.34)	5.008** (0.34)
Duration 4 years	4.873** (0.34)	4.766** (0.34)
Duration 5 years	4.672** (0.35)	4.562** (0.34)
Duration 6 years	4.630** (0.35)	4.520** (0.35)
Duration 7 years	4.634** (0.36)	4.519** (0.36)
Duration 8 years	4.355** (0.38)	4.235** (0.38)
Duration 9 years	4.453** (0.41)	4.324** (0.41)
AGE	0.065 (0.05)	0.085 (0.05)
AGE SQUARED	- 0.0007 (0.0005)	- 0.0009* (0.0005)
Female head	0.039 (0.06)	0.056 (0.06)
Member second pillar	- 0.478** (0.05)	- 0.439** (0.05)
	Education level head	
Low secondary general	0.101 (0.08)	0.104 (0.08)
Low secondary technical	- 0.167** (0.08)	- 0.150** (0.08)
Low secondary professional	- 0.039 (0.08)	- 0.027 (0.07)
High secondary general	- 0.026 (0.08)	- 0.131 (0.08)
High secondary technical	- 0.120 (0.10)	- 0.103 (0.09)
High secondary professional	0.170 (0.11)	0.149 (0.11)

University	- 0.351** (0.07)	- 0.312** (0.07)
Education level spouse		
Low secondary general	0.280** (0.08)	0.252** (0.07)
Low secondary technical	- 0.041 (0.14)	- 0.037 (0.13)
Low secondary professional	- 0.006 (0.08)	- 0.006 (0.08)
High secondary general	0.080 (0.12)	0.080 (0.12)
High secondary technical	0.005 (0.16)	0.025 (0.16)
High secondary professional	0.192 (0.12)	0.173 (0.12)
University	- 0.366** (0.11)	- 0.335** (0.11)
Covariance parameter estimates	ρ_1=- 0.59 (0.04)	ρ_1=- 0.51 (0.02)
Formal likelihood ratio test of significance of covariance parameter estimates:H_0: $\rho_1 = 0$		
	χ_1^2 =9987 (Pr> χ_1^2 :<0.0001)	χ_1^2 =13024 (Pr> χ_1^2 :<0.0001)
, 2loglikelihood	459184	446876
Residual	0.97 (0.006)	0.97 (0.006)
Number observations	72844	72844
* denotes significance at 5% level, ** denotes significance at 1% level; standard errors in parenthesis. The reference person is a male head with primary school education, is no member of second pillar, with a spouse with primary school education, in 1992, living in province of Antwerp. Year and province dummies are included in all equations.		

As one can see, the coefficients and standard errors are very similar. The predicted hazard rates are barely distinguishable and therefore not plotted. This means that we should not worry to much about having used a logistic distribution instead of a extreme value distribution for the random error terms. Finally, it is sometimes argued that the extreme value may be preferred to the logistic distribution for theoretical reasons since the former is the discrete-time equivalent of the continuous proportional hazard model: “The proportional hazard model is often regarded to be useful as reduced-form model for duration analysis. The resulting estimates are generally interpreted with the help of some economic theory. However the proportional hazard model specification is not derived from economic theory and it remains to be seen whether the proportional hazard specification is actually able to capture important theoretical justifications and conversely whether the proportional hazard specification can be generated by theory”. In particular, “first, the proportionality restriction of the (M)PH model can in general not be justified on economic theoretical grounds. Second if the optimal strategy is myopic (because of repeated search or the discount rate is infinite) then this restriction often follows from economic theory”.[61]

[61] van den Berg (2001), p.25, 29.

Appendix IV: Sensitivity of the results to the inclusion of the self-employed

We interpreted the strong empirical support for true duration dependence as related to work disincentives, stigma or depreciation of human capital. One might argue these arguments concern primarily those covered by the social security system of the employed while the mechanism behind the observed poverty among the self-employed may in addition be related to their non-declaration of incomes. To check whether the latter mechanism would dominate our results, we re-estimated exit and re-entry rates simultaneously while controlling for observed and unobserved heterogeneity and without left-censored spells but did this separately, in table 9, for those who have been employed and those who have been self-employed.

Table 9: Multiple-spell discrete-time hazard model with unobserved effects for those covered by social security system of the employed and self-employed[62]

Variable	Employed	Self, employed
Exit		
Intercept	- 16.414** (1.60)	- 17.77** (3.16)
Duration 1 year	6.223** (0.29)	7.192** (1.00)
Duration 2 years	5.351** (0.29)	6.442** (1.00)
Duration 3 years	5.124** (0.29)	6.227** (1.01)
Duration 4 years	5.150** (0.30)	6.317** (1.01)
Duration 5 years	4.969** (0.30)	6.106** (1.02)
Duration 6 years	4.858** (0.31)	5.797** (1.03)
Duration 7 years	4.656** (0.33)	5.748** (1.03)
Duration 8 years	4.711** (0.34)	4.806** (1.09)
Duration 9 years	4.501** (0.38)	6.007** (1.06)
AGE	0.355** (0.05)	0.404** (0.09)
AGE squared	- 0.003** (0.0004)	- 0.004** (0.0008)
Female head	- 0.589** (0.005)	- 0.141 (0.122)

[62] Since the sample consists of repeated observations on the same household, standard errors are adjusted to account for the dependence at the level of the household.

Member second pillar	0.372** (0.04)	0.170** (0.09)
Education level head		
Low secondary general	0.122 (0.08)	0.036 (0.119)
Low secondary technical	0.305** (0.08)	0.312** (0.123)
Low secondary professional	0.114 (0.08)	0.370** (0.111)
High secondary general	0.304** (0.10)	0.176 (0.136)
High secondary technical	0.423** (0.10)	0.187 (0.146)
High secondary professional	0.249* (0.125)	- 0.128 (0.175)
University	0.367** (0.07)	0.542** (0.113)
Education level spouse		
Low secondary general	- 0.106 (0.08)	0.042 (0.11)
Low secondary technical	- 0.031 (0.14)	0.006 (0.17)
Low secondary professional	0.062 (0.08)	- 0.010 (0.13)
High secondary general	- 0.082 (0.12)	0.526** (0.16)
High secondary technical	0.137 (0.17)	- 0.041 (0.23)
High secondary professional	0.013 (0.12)	- 0.035 (0.20)
University	0.213* (0.12)	0.394** (0.14)
Re-entry		
intercept	- 8.691** (2.20)	- 7.703** (3.23)
Duration 1 year	5.610** (0.35)	7.196** (1.00)
Duration 2 years	5.033** (0.36)	6.552** (1.00)
Duration 3 years	4.845** (0.36)	6.201** (1.01)
Duration 4 years	4.654** (0.37)	5.823** (1.01)
Duration 5 years	4.507** (0.37)	5.476** (1.02)
Duration 6 years	4.433** (0.38)	5.543** (1.03)
Duration 7 years	4.289** (0.40)	5.785** (1.03)
Duration 8 years	4.270** (0.42)	4.909** (1.08)
Duration 9 years	4.275** (0.46)	5.409** (1.11)
AGE	0.114* (0.07)	- 0.019 (0.10)

AGE SQUARED	- 0.001** (0.0006)	0.0003 (0.0008)
Female head	0.210** (0.07)	- 0.162 (0.13)
Member second pillar	- 0.486** (0.06)	- 0.068 (0.10)
Education level head		
Low secondary general	- 0.190 (0.11)	- 0.051 (0.13)
Low secondary technical	- 0.324** (0.10)	- 0.258* (0.13)
Low secondary professional	- 0.287** (0.10)	- 0.141 (0.115)
High secondary general	- 0.250** (0.12)	- 0.520** (0.14)
High secondary technical	- 0.257** (0.12)	- 0.267 (0.17)
High secondary professional	- 0.056 (0.15)	- 0.037 (0.16)
university	- 0.674** (0.09)	- 0.460** (0.11)
Education level spouse		
Low secondary general	0.230* (0.10)	0.214* (0.11)
Low secondary technical	- 0.106 (0.18)	0.007 (0.20)
Low secondary professional	- 0.133 (0.11)	0.136 (0.12)
High secondary general	0.129 (0.15)	- 0.276 (0.20)
High secondary technical	0.002 (0.20)	- 0.162 (0.30)
High secondary professional	0.209 (0.15)	0.080 (0.22)
university	- 0.406** (0.15)	- 0.693** (0.16)
Covariance parameter estimates	ρ_1 =- 0.624 (0.05)	ρ_1 =- 0.34 (0.06)
Formal likelihood ratio test of significance of covariance estimate: H_0: $\rho_1 = 0$		
	χ_1^2 =7644 (Pr> χ_1^2 : <0.0001)	χ_1^2 =3384 (Pr> χ_1^2 : <0.0001)
, 2loglikelihood	351538	113860
residual	0.94 (0.006)	1.107 (0.01)
Number observations	55541	17303

* denotes significance at 5% level, ** denotes significance at 1% level; standard errors in parenthesis. The reference person is a male head with primary school education, is no member of second pillar, with a spouse with primary school education, in 1991, living in province of Antwerp. Year and province dummies were included in all equations.

Since it may be difficult to interpret the results of a nonlinear model, we plot the predicted hazard rates.

Figure 7: Predicted exit rates for those covered as employed and self, employed

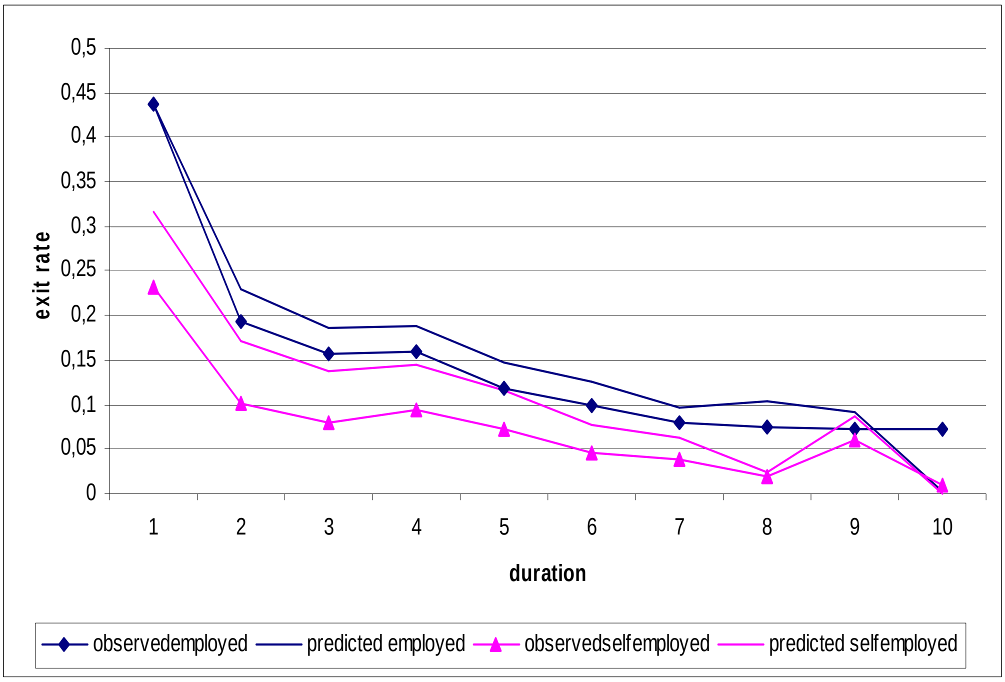
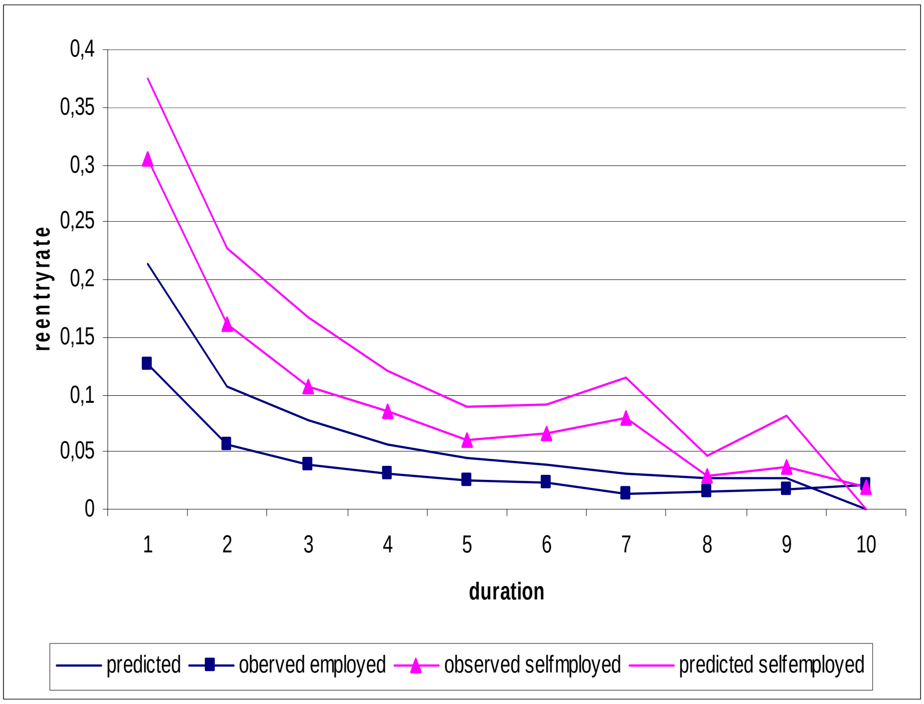


Figure 8: Predicted re-entry rates for those covered as employed and self, employed



The exit rate out of poverty is higher for employed than self-employed and the re-entry rate into poverty is lower among employed than self-employed. This would mean self-employed are more likely to be persistently poor. It is, however, well-known that self-employed do not declare all their income and this may play a role in explaining this pattern[63]. We do not study this further, since what is important is that for the employed, after controlling for observed and unobserved characteristics, it remains true that negative dependence in poverty is a true phenomenon. The fact that the pattern of exit and re-entry rates remains downward sloping, also after controlling for observed and unobserved characteristics is a first reason why we did not treat these groups as different in the main analysis. In addition, about 75% of the self-employed have a mixed career as self-employed and employed and thus benefit social security rights, especially old-age pension rights[64], in both systems. It would be necessarily arbitrary to split the population in two groups and it would result in a loss of information.

[63] "Tax evasion and fiscal fraud attain in Belgium significant levels" (Franck (1987)). See also HUB, Research paper 2008, 19.

[64] National Office for Pensions (2008), "Jaarlijkse statistiek van de uitkeringsgerechtigden: toestand op 1 januari 2007", Brussels, p.24.