The author would like to make the following comment:

This paper is a preliminary draft and not the final version. Section 5, investigating the long-run determinants of preference heterogeneity in labor market transitions, is work-in-progress. A final draft of the paper will be made available to the "Crete 2011" Committee by May 2, 2011.

The author assumes full responsibility of any errors.
Any remarks are welcome.

*Thank you.*_

Employment transitions of older persons in Britain: State dependence and long-run determinants

By Elpianna Emmanouilidi[1 2]

Athens University of Economics and Business (AUEB)
Economics Deparment

Preliminary draft
Submission date: February.14^{th}, 2011

This paper uses a balanced panel constructed on the basis of the first four waves (1998-2001) of the English Longitudinal Survey of Ageing (ELSA) in order to study the employment transitions of older persons (the 50-65 age group) in Britain. Several mutually exclusive employment states are considered ranging from activity to inactivity: full-time employment, part-time employment, self-employment, permanently sick or disabled and/or fully retired, homemakers, non-participation. The underlying structural model is a first-order Markov process assuming perfect mobility and preference heterogeneity among states. The structural assumptions are tested by means of conditional likelihood estimation of a dynamic multinomial logit model with unobserved heterogeneity. There is evidence of true state dependence and duration dependence in employment choices. *Long-run determinants of individual preferences are studied next with emphasis on health status, socioeconomic status and work experience (work-in-progress).*

Keywords and phrases: xxx.

[1]Author contact details: Doctoral Studies Program, Economics Department, Athens University of Economics and Business, 47 Evelpidon & 33 Lefkados, 7th Floor, Room 705, Athens 113 62, Greece. Email: elpemman@aueb.gr.

[2]The author is particularly grateful to Thierry Magnac for providing access to his dynamic fixed effects logit Gauss code as well as to Ekaterini Kyriazidou for useful remarks and excellent PhD thesis supervision. This research is supported by the 2010 Basic Research Funding Program (BRFP) of the Athens University of Economics and Business (AUEB), Athens, Greece.

1 Introduction

The populations of most developed economies are ageing and exhibit increased life expectancy. As a result, these economies have experienced a sharp decline in the participation of older persons in the labor force with adverse budgetary consequences mainly because their social security systems are financed on a pay-as-you-go (PAYG) basis. The mutuality principle of the PAYG institution has proved financially vulnerable to the ageing of the populations since the proportion of the population that has retired and is collecting benefits exceeds the fraction that remains active in the labor force and is paying for these benefits. An increase in the ratio of retiring-to-active workers is expected to significantly reduce the productive labor potential of a developed economy and slow down its economic growth. These developments have triggered the reform of social security systems with a direct effect on labor market policy (see [25], [30]). Moreover, they justify the increased attention on the part of economic policymakers on the retirement behavior of older persons, in particular of the 50-65 age group. Thus, in an attempt to prevent a potentially massive exit from the labor market, financial and other institutional incentives have been designed in order to encourage the 50-65 target group to remain active. For instance, in the UK, the age for state pension eligibility is set to increase from 65 years for men and 60 years for women to 68 years for both by 2046, while recent European Union Council proposals suggest an increase in the normal retirement age to 67 years of age for both men and women across Europe.

However, institutional intervention will never be fully efficient unless it internalizes the fact that individual labor force participation is an intertemporal process characterized by individual heterogeneity. This diversity may be explained both by observed factors, such as health and private wealth status, as well as unobserved factors, mainly individual preferences for work, as opposed to non-participation, and institutional barriers to entry/exit among employment states. For instance, a sudden deterioration in individual health status may trigger a change of individual labor status irrespective of existing supply-side policies. Or, institutional barriers governing employment contracts may influence the duration of activity/inactivity spells. Thus, controlling for both observed and unobserved individual heterogeneity, the decision to remain active in the labor market is expected to exhibit a certain degree of persistence in the form of state and duration dependence.

This paper uses a dynamic panel approach in order to study the employment dynamics of older persons in Britain (the 50-65 age group) in a discrete-time discrete-choice setting and identify both observable and unobservable factors which might influence their participation in the labor force and particularly, their retirement behavior. Transitions in and out of continued work and, in particular, exit rates, of the 50-65 age group are studied in a multivariate context with employment states ranging from full-time employment to non-participation: full-time employment, part-time employment, self-employment, permanently sick or disabled and/or fully retired, homemakers, non-participation. The presence (or absence) of any persistence in the employment choices of the target group is of particular interest as it would signal the existence (or absence) of a pattern of individual responses to labor market conditions. In general, individual behavior is expected to diverge from group average behavior as it reflects unobserved heterogeneity in choices. Following this line of reasoning, it would be useful to determine the presence (or absence) of such persistence in individual employment choices in the first place, and then investigate the sources of preference heterogeneity considering a range of long-run determinants, such as socio-economic characteristics, financial incentives, private wealth and health status, with emphasis on health status, socioeconomic status and work experience.

The research sample is a balanced panel consisting of individual employment histories created on the basis of employment and pension information from waves 0-4 of the English Longitudinal Study of Ageing (ELSA). The empirical results provide evidence in favor of persistence in individual labour market transitions in the form of state and duration dependence conditional on unobserved preference heterogeneity. The long-run determinants of individual preferences are studied next *(work-in-progress)*.

2 Recent empirical evidence

So far, the bulk of the empirical literature on health and retirement of older persons has focused on the determinants of early retirement. This paper brings together two strands of literature, one focusing on the retirement behavior of older persons and its determinants and the other examining the employment transitions of the working population focusing mainly on younger age groups.

2.1 Determinants of early retirement

Even though financial incentives, such as the option value of work, are considered to be the prevalent determinant of labor force exit of older persons in the empirical literature (for example, [51], [25], [30], [10] and [6]) a growing number of studies recognize the importance of individual-specific factors, in particular health (for example, [14], [23], [46], [24], [42], [15] and [29]), private wealth (for example, [9]) and socioeconomic status as co-determinants of the labor force exit of older persons.

In addition, institutions and social security policy are considered as well (for example, [13] and [5]). Of interest to this paper, [5] uses data from the English Longitudinal Survey of Ageing (ELSA) and finds evidence against the legislation of a uniform normal retirement age in England since retirement is a function of individual health status hence individual-specific.

Standard economic theory underlying the relationship between retirement and health (see [38]) suggests that, other things being equal, poorer health causes negative selection in continued work for several reasons: it may raise the current disutility of work, it may reduce the return from work if there is a relationship between poor health and low wages, or it may entitle the individual to non-wage income, such as disability benefits, contingent on non-activity. For instance, [7] use a sample of 2497 men with complete employment histories on the basis of the U.S. National Longitudinal Survey (NLS) of mature men for the period 1966-1982 and find that bad health, along with age and lack of education, increase the probability of retirement. [23] agree that there is a health effect on labor force participation but point out that there is a lack of consensus on its magnitude and its relative importance compared to other determinants. [47] finds that health shocks experienced by Germans aged 40-59 increase the probability of unemployment by 84% and of labor force exit by 200%. *[49]* finds that individuals over 50 years who suffer a health shock experience an immediate 15% decrease in the probability of continued work. This effect is not stationary but decreases to 1/3 of its original magnitude five years after the shock.

2.2 Employment transitions

Most empirical studies analyzing labor market transitions explain individual participation behavior on the basis of health status, among others. This is so because health status is usually treated as a strictly exogenous regressor in the retirement equation unless a structural approach is adopted by the researcher (for example, [15]). Moreover, most evidence is derived on the basis of cross-sectional data which severely limits the ability to study unobserved heterogeneity in individual preferences. Finally, the majority of the studies using longitudinal data focus on a single hazard at a time, usually the transition out of continued work and into retirement (and vice versa). This paper is concerned with the employment dynamics of older persons in a dynamic multivariate context and aims to make a valuable contribution in the aforementioned directions.

[24] examine the effect of ill health on the retirement decision of older people in Britain using BHPS data for the period $1991-1998$. They find that adverse health shocks predict individual retirement behavior of older persons, that health shocks have symmetric effects on transitions in and out of economic activity, and that both current and lagged individual health stock changes affect current transitions in and out of economic activity. These results imply some form of state dependence in individual activity on the basis of health status. The part of the study concerning individual retirement behavior uses conditional maximum likelihood in order to estimate a binary logit model with individual fixed effects and strictly exogenous regressors (see [16]).

[29] analyze the role of the health status of the working-age population on employment exits/entries using data from the first twelve waves of the British Household Panel Survey $(1991-2002)$. They use a discrete-time duration model to estimate the effect of health on the hazard of becoming non-employed as well as on the hazard of becoming employed. Their empirical findings show that individual health is an important determinant of employment transitions and that health effects are greater for men. However, the magnitude of the health effects depends on the measure of health used.

Larger effects of health on labor outcomes are found for prime-age individuals (individuals in their 30-40s) measured on the basis of wage loss on the onset of a permanent illness. Using UK longitudinal data, [22] find that reduced psychological health reduces the hourly wage for men, while excellent self-assessed health increases the hourly wage for women. [28] use Spanish data in order to analyse the effects of a health shock on the probability of transiting from employment into unemployment or inactivity. They find that a health shock decreases the probability of remaining in employment by 5% and increases the probability of transiting into inactivity by 3.5%.

Relatively fewer studies analyse the effect of ill-health on transitions into employment (entries rather than exits). Existing evidence suggests that individuals with impaired health have longer unemployment spells In particular, [50] uses Canadian data to examine the effects of health limitations on the type and amount of activity that individuals can undertake at work. He finds that individuals with impaired health experience longer unemployment spells because they have a lower chance of leaving unemployment. [12] use calendar data from the first seven waves of the BHPS to analyse transitions from unemployment into part-time work, self-employment and economic inactivity. They find that a limiting health condition preceding an unemployment spell doubles the exit rate into economic inactivity.

The number of studies analysing the effect of health on labor market transitions for younger individuals is less extensive (see [29]). Of particular interest to the present paper is the study of [41] who uses the $1990-1992$ wave of the French Labor Force Survey to study the dynamics of labor market transitions of young people in a purely dynamic multivariate context, ranging from non-participation to stable employment with temporary employment and paid training as intermediate employment states. He finds that controlling for unobserved heterogeneity refutes his original finding according to which temporary employment and paid training were ladders to stable employment. Furtheron, he uses a maximum simulated likelihood framework in order to find the long-run determinants of unobserved heterogeneity, that is of individual preferences for work, considering health status along with other factors.

Finally, [8] use monthly data from waves $1-11$ of the German Socio-Economic Panel (GSEOP) in order to study the retirement behavior of older married couples based on a discrete-time competing hazards model of labor force transitions relying on joint spouse employment status. They find that the probability of one spouse exiting employment is much larger if the other spouse is not employed than employed. Similarly, one member of a couple is much more likely to enter employment if the spouse is employed than not employed.

3 Data description

The English Longitudinal Study of Ageing (ELSA) is a biennial survey on ageing of people aged 50 years old and over including their younger partners (aged 50 years old or less) living in private households in England. It is designed on the basis of the US Health and Retirement Survey (HRS).

ELSA is quite appealing for the purposes of the present research because it is a relatively unexplored panel with a long enough time dimension, representative of the recent developments on ageing, health and retirement. As a matter of fact, the population of Britain is getting older as people are living longer. For instance, [43] find that as many as 1/3 of individuals in Britain are aged 50 years old and over and account for 1/2 of national spending and 3/4 of national wealth.

The research sample used in this paper consists of data from five consecutive waves: HSE (Health Survey of England) data from wave 0 ($1998-1999, 2001$) and data from ELSA biennial waves $1-4$ ($2002-2009$). Wave 0 observations have been left-censored to 2001 implying that 2001 is the starting period. The sample selection criterion consists of including those respondents who have had a productive interview in all waves considered, aged 50-65 years old in 2001. As a consequence, the research sample is a balanced panel consisting of 3690 (out of 6232) respondents born during $1936-1951$ of which 1660 males and 2030 females.[3] 1799 of the panel respondents live in households with > 1 members while the remaining 1891 live in single-member households. As far the time dimension of the panel is concerned, there are 9 yearly or 18 half-yearly (double frequency) or 36 quarterly ($4\times$ frequency) observations. Individual respondents have been located on the time axis (equal time intervals) on the basis of individual interview dates (unequal time intervals).

Individual employment states have been deterined on the basis of answers provided to questions included in the ELSA 'Employment and Pensions' module for each consecutive wave. A total of $M = 5$ mutually exclusive employment states have been considered: (a) full employment ($EMPL$), self-employment ($SELF$), permanently sick (or disabled) or retired ($PSDRET$), homemakers ($HOME$), $OTHER$ (indicating non-participation) in case of yearly and half-yearly observations (b) full employment ($EMPLFT$), part-time employment ($EMPLPT$), self-employment ($SELF$), permanently sick (or disabled) or retired or homemakers ($PSDRETHO$), $OTHER$ (indicating non-participation) in case of quarterly observations. The last category, $OTHER$, is the reference state for identification purposes. Since self-reported answers are subject to reporting error, more objective information, such as the receipt of a disability benefit, have been considered in order to classify respondents into employment states.

4 Employment dynamics

The structural model underlying labor market transitions of older persons is a purely dynamic, multivariate Markov transition process with preference heterogeneity and no barriers to entry/exit among labor market states (see [41] and

[3] The full sample without a year of birth selection rule is a balanced panel consisting of 6232 respondents born in the interval $1932-1952$ (and aged $49-69$ years old in 2001) of which 2655 males and 3577 females.

[44]; see also [26] and [21]). The absence of barriers implies that the rates of arrivals of job offers from any other labor market state are equal across states. The structural assumptions are tested on the basis of estimation of a first-order Markov transition process conditional on unobserved heterogeneity.

The estimation framework consists of the following steps. Initially, a homogeneous first-order Markov transition model without heterogeneity is estimated conditional on the vector of initial conditions in order to describe the employment dynamics of the panel respondents in the period $2001-2009$. Next, two dynamic multinomial models are estimated: the first one (the model under the null) with observed heterogeneity modelled as a linear index of time-invariant regressors, and the second one (the model under the alternative) conditional on unobserved heterogeneity, that is with fixed effects. Estimation results are used in order to test the null hypothesis of no unobserved heterogeneity in individual preferences. A rejection of the null would indicate the presence of true state dependence in individual employment choices.

4.1 Conditional estimation of a homogeneous first-order Markov transition model

Employment transitions among the M states during the sample period are estimated on the basis of [4] (see [4] and [1]). An individual's employment history is treated as a Markov chain consisting of a sequence of states $\{j(0), j(1), ..., j(T)\}$. Given the initial state $i(0)$, there are M^T possible sequences of mutually exclusive events with occurence probabilities $P_{j(0)j(1)}...P_{j(T-1)j(T)}$. The transition process is assumed to be of the first order implying that it has one-period memory. Moreover, it is both homogeneous and stationary, that is $P^i_{jk}(t) = P_{jk}$ $\forall i$ and $\forall t$ respectively, where $P_{jk}(t)$ denotes the transition probability to destination state k at time t given origin state j at time $t-1$. Assuming that $n_j(0)$ i.e. the number of individuals in state j at time 0, is nonrandom, and that $n_{j(0)j(1)...j(T)}$ is the number of individuals whose observed sequence is $\{j(0), j(1), ..., j(T)\}$, define $\{n_{jk}(t)\}$ as the set of M^2T sufficient statistics for the observed sequences where n_{jk} denotes the number of transitions (exits) from origin state j to destination state k. Assuming $\sum_{k=1}^{M} n_{jk}(t)$ observations on a multinomial distribution, the density of $n_{jk}(t)$ *conditional on* $n_j(0)$ is

$$\prod_{t=1}^{T}\left\{\prod_{j=1}^{M}\left[\frac{\sum_{k=1}^{M} n_{jk}(t)!}{\prod\limits_{k=1}^{M} n_{jk}(t)!}\prod_{k=1}^{M} P_{jk}(t)^{n_{jk}(t)}\right]\right\}$$

Since the process is stationary, the conditional density of a given ordered set of sequences for all N individuals (describing all sequences in the NMT space) can be simplified, as follows:

$$
\begin{aligned}
L &= \prod \left[P_{j(0)j(1)}\left(1\right)...P_{j(T-1)j(T)}\left(T\right)\right]^{n_{j(0)j(1)\dots j(T)}} \\
&= \prod_{t=1}^{T} \prod_{j(t-1),j(t)} P_{j(t-1)j(t)}\left(t\right)^{\sum n_{j(0)j(1)\dots j(T)}} \\
&= \prod_{t=1}^{T} \prod_{j(t-1),j(t)} P_{j(t-1),j(t)}\left(t\right)^{n_{j(t-1)j(t)}(t)} \\
&= \prod_{k}\prod_{j} {P_{jk}}^{n_{jk}} \qquad \text{where } P_{jk}\left(t\right) = P_{jk} \ \forall t
\end{aligned} \tag{1}
$$

The *conditional density* L is the same, apart for a factor that does not depend on P_{jk}, as that obtained for M independent samples where the j^{th} sample consists of $n_j^* = \sum_k n_{jk}$ multinomial trials with probabilities P_{jk}. The transition probability estimates P_{jk} are obtained by maximizing the conditional density 1 with respect to P_{jk} subject to M^2 non-negativity constraints $P_{jk} \geq 0$ and M equality constraints $\sum_{k=1}^{M} P_{jk} = 1, j = 1, ..., M$. The first order condition of the constrained Lagrangean yields an expression for n_{jk} :

$$
S = \sum_{k}\sum_{j} n_{jk} \log P_{jk} - \sum_{j} \lambda_j \left(\sum_{k=1}^{M} P_{jk} - 1\right)
$$

$$
n_{jk} = \lambda_j P_{jk}
$$

Summing over k (and using the equality constraints) yields the maximum likelihood estimate (MLE) of the transition parameters

$$
\hat{P}_{jk} = \frac{n_{jk}}{\sum_k n_{jk}} = \frac{n_{jk}}{n_j^*} = \frac{\sum\limits_{t=1}^{T} n_{jk}\left(t\right)}{\sum\limits_{l=1}^{M}\sum\limits_{t=1}^{T} n_{jl}\left(t\right)} = \frac{\sum\limits_{t=1}^{T} n_{jk}\left(t\right)}{\sum\limits_{t=0}^{T-1} n_{j}\left(t\right)} \tag{2}
$$

It can be proved (see [4]) that the variables $N^{1/2}\left(\hat{P}_{jk} - P_{jk}\right)$ have a joint normal limiting distribution with means 0 and variances $\frac{1}{\phi_j} P_{jk}\left(1 - P_{jk}\right)$ where

$$
\phi_j = \sum_{l=1}^{M}\sum_{t=1}^{T} \eta_l P_{lj}^{[t-1]}
$$

$$
0 < \eta_l = p \lim_{\sum n_l(0) \to \infty} \frac{n_l\left(0\right)}{\sum n_l\left(0\right)}
$$

and

$$
\hat{P}_{jk}^{[t]} = \sum_{l=1}^{M} n_{l;jk}\left(t\right) / \sum_{l}\sum_{k} n_{l;jk}\left(t\right)
$$

A summary of employment transitions in the period $2001-2009$ is provided in *Tables 1a-c* and *2a-c*. *Tables 1a-c* show state marginal probabilities over the period and mean duration estimates. The estimated average duration in state j is defined as

$$\hat{d}_j = \frac{T_j}{n_j} \tag{3}$$

where T_j is the sum of durations of all spells of type i and n_j is the number of exits from state j (excluding right-censoring). *Tables 2a-c* are two-way $M \times M$ tables of transition probability estimates $\hat{P}_{jk}$ (whose rowsums equal one).

Table 1a-b.

Descriptive statistics for the period $2001 - 2009$ yearly.
Sample size: $N = 3690$.

	Mean state duration $\hat{d}_j$ (time periods)	Marginal probability (%)	#Obs.
EMPL	6	59	2166
SELF	1	9	325
PSDRET	5	23	848
HOME	1	5	174
OTHER	1	5	178

Table 1b.

Descriptive statistics for the period $2001 - 2009$ half-yearly.
Sample size: $N = 3690$.

	Mean state duration $\hat{d}_j$ (time periods)	Marginal probability (%)	#Obs.
EMPL	12	59	2185
SELF	3	9	323
PSDRET	9	22	804
HOME	2	5	167
OTHER	1	6	211

Table 1c.

Descriptive statistics for the period $2001 - 2009$ quarterly.
Sample size: $N = 3690$.

	Mean state duration $\hat{d}_i$ (time periods)	Marginal probability (%)	#Obs.
EMPLFT	26	53	1946
EMPLPT	5	7	248
SELF	4	9	322
PSDRETHO	17	26	950
OTHER	2	6	224

Table 2a.

% Transition parameters $\hat{P}_{jk}$ for the period $2001 - 2009$ yearly.

Origin ↓ Destination →	EMPL	SELF	PSDRET	HOME	OTHER
EMPL	0	14 (.036)	60 (.052)	13 (.036)	13 (.036)
SELF	50 (.042)	0	33 (.039)	6 (.020)	11 (.026)
PSDRET	12 (.010)	3 (.005)	0	69 (.014)	16 (.011)
HOME	7 (.007)	1 (.003)	89 (.009)	0	3 (.004)
OTHER	12 (.010)	4 (.006)	71 (.014)	13 (.011)	0

All estimates are statistically significant at the 1% level.

Table 2b.

% Transition parameters $\hat{P}_{jk}$ for the period $2001 - 2009$ half-yearly.

Origin ↓ Destination →	EMPL	SELF	PSDRET	HOME	OTHER
EMPL	0	15 (.033)	59 (.045)	13 (.031)	13 (.031)
SELF	50 (.034)	0	33 (.032)	6 (.016)	11 (.021)
PSDRET	12 (.010)	3 (.006)	0	69 (.014)	16 (.011)
HOME	7 (.007)	1 (.003)	89 (.009)	0	3 (.004)
OTHER	22 (.014)	8 (.009)	59 (.016)	11 (.010)	0

All estimates are statistically significant at the 1% level.

Table 2c.

% Transition parameters $\hat{P}_{jk}$ for the period $2001 - 2009$ quarterly.

Origin ↓ Destination →	EMPLFT	EMPLPT	SELF	PSDRETHO	OTHER
EMPLFT	0	14 (.023)	15 (.023)	61 (.032)	11 (.021)
EMPLPT	73 (.025)	0	2 (.008)	21 (.023)	4 (.011)
SELF	49 (.022)	1 (.005)	0	39 (.021)	11 (.013)
PSDRETHO	45 (.012)	1 (.003)	11 (.008)	0	43 (.012)
OTHER	21 (.013)	5 (.007)	9 (.009)	65 (.015)	0

All estimates are statistically significant at the 1% level.

Employment ($EMPL$ in case of yearly and half-yearly data, and $EMPLFT$ + $EMPLPT$ in case of quarterly data) and retirement ($PSDRET$ + $HOME$ in case of yearly and half-yearly data and $PSDRETHO$ in case of quarterly data) are the dominant employment states over the period (see *Tables 1a-c*): 59% $EMPL$ and 23% $PSDRET$ in case of yearly data, 59% $EMPL$ and 22% $PSDRET$ in case of half-yearly data, and (53% $EMPLFT$ + 7% $EMPLPT$

=)60% $EMPL$ and 26% $PSDRETHO$ in case of quarterly data. Around 12% (= [7/ (53 + 7)] %) of the employed are part-timers (see *Table 1c*).

As far as transition estimates are concerned, retirement ($PSDRET$ +/ $HOME$ in case of yearly and half-yearly data and $PSDRETHO$ in case of quarterly data) is the main exit route for all employment states apart from self-employment and part-time employment in case of quarterly data. The main exit route of these two states is full-time employment (see *Tables 2a-c*). Actually, the transition from part-time to full-time employment (see *Table 2c*) carries the coefficient estimate with the highest value in the quarterly transition matrix, 73%. Moreover, half of the self-employed transit into full-time employment over the period on a quarterly basis. It seems that self-employment and part-time employment (in case of higher-frequency data) are intermediate states to retirement. It remains to be seen whether this finding will still be valid when individual preference heterogeneity will be controlled for.

4.2 Dynamic multinomial logit estimation

At this point, the index of states will be modified so as to include a zero indicating the reference state i.e. non-participation. Thus, there will be a total of $M+1$ states including the reference state with $m = 0, 1, ..., M$. Let the latent variable y^*_{ikt} describe the propensity of individual i being in state k at time t :

$$y^*_{ikt} = \sum_{j=0}^{M} \delta_{jk} 1 \{y_{i,t-1} = j\} + \epsilon_{ikt} \tag{4a}$$

$$\epsilon_{ikt} = \alpha_{ik} + u_{ikt} \tag{4b}$$

Current-state individual propensity is a function of lagged state variables, y_{it}, and of an additive error term, ϵ_{ikt}, which includes individual heterogeneity, α_{ik}. In the present context, the α_{ik} describe individual preferences for paid work as a function of the current labor market state. The observed state has maximum marginal propensity:

$$y_{it} = k \text{ if } y^*_{ikt} = \max_j \left(y^*_{ijt}\right) \tag{5}$$

Assuming that the errors u_{ikt} conditional on α_{ik}, $u_{ikt}|\,\alpha_{ik}$, are Type II extreme value distributed and independent across states, individuals and time periods, the probability of individual i being in state k at time t conditional on having been in state j at time $t-1$ is given by

$$\begin{aligned} &\Pr\left(y_{it} = k|\, y_{i,t-1} = j; \delta, \alpha\right) = \\ &= \frac{\exp\left(\delta_{jk} - \delta_{j0} + \alpha_{ik} - \alpha_{i0}\right)}{1 + \sum_{l \neq 0} \exp\left(\delta_{jl} - \delta_{j0} + \alpha_{il} - \alpha_{i0}\right)} = \\ &= \frac{\exp\left(\delta_{jk} + \alpha_{ik}\right)}{1 + \sum_{l \neq 0} \exp\left(\delta_{jl} + \alpha_{il}\right)}; \qquad \delta_{j0} = \alpha_{i0} = 0 \end{aligned} \tag{6}$$

The probability in (6) specifies a dynamic multinomial discrete-choice model with unobserved heterogeneity. Reference state parameters are not identifiable and are thus set to zero for estimation purposes. The odds of being in state k (relative to state 0) conditional on any lagged state $j \neq k$ is a function of individual heterogeneity associated with that state, α_{ik} :

$$\frac{\Pr\left(y_{it}=k|\, y_{i,t-1}=j;\delta,\alpha\right)}{\Pr\left(y_{it}=0|\, y_{i,t-1}=j;\delta,\alpha\right)}=\exp\left(\delta_{jk}+\alpha_{ik}\right)$$

Individual state heterogeneity may be removed from the odds ratio, by reexpressing it as follows:

$$\frac{\frac{\Pr(y_{it}=k|y_{i,t-1}=j;\delta,\alpha)}{\Pr(y_{it}=0|y_{i,t-1}=j;\delta,\alpha)}}{\frac{\Pr(y_{it}=k|y_{i,t-1}=0;\delta,\alpha)}{\Pr(y_{it}=0|y_{i,t-1}=0;\delta,\alpha)}}=\exp\left(\delta_{jk}-\delta_{0k}\right)=\exp\delta_{jk};\qquad \delta_{0k}=0$$

Or, in terms of log-retention rates:

$$\log\left[\frac{\frac{\Pr(y_{it}=j|y_{i,t-1}=k;\delta,\alpha)}{\Pr(y_{it}=0|y_{i,t-1}=k;\delta,\alpha)}}{\frac{\Pr(y_{it}=j|y_{i,t-1}=0;\delta,\alpha)}{\Pr(y_{it}=0|y_{i,t-1}=0;\delta,\alpha)}}\right]=\delta_{kj} \tag{7}$$

For instance, a positive rate δ_{kj} would imply that the odds of being in an employment state (compared to non-participation) are larger when the origin state is some other employment state rather than if it were non-participation (the reference state) implying that non-participation is associated with relatively greater delays of reentering the labor market.

4.2.1 With observed heterogeneity

This section presents estimation results of the model in (6) under the null hypothesis of no unobserved state heterogeneity in individual preferences for work. Following [41], individual state heterogeneity α_{ik} is assumed to be controlled for by a set of observable and time-invariant individual characteristics. If the set of controls consisted of time-varying regressors, the estimation framework of [35] would be appropriate instead. In particular, α_{ik} is specified as a linear index of time-invariant discrete variables, such as sex, highest educational qualification, social class and the level of physical activity,

$$\alpha_{ik}=X_i\beta_k$$

and so the original current-state individual propensity (**??**) becomes

$$y^*_{ikt}=\sum_{j=0}^{M}\delta_{jk}1\left\{y_{i,t-1}=j\right\}+X_i\beta_k+u_{ikt} \tag{8}$$

The variables on the right-hand side of (8) include state dummies describing the first lag of the dependent variable and time-invariant controls for observed state heterogeneity. As a result, the linear control index is *alternative-varying* *but time-invariant* while the lagged state coefficients are both *alternative- and time-varying.* Removing heterogeneity in this way is similar to ignoring index i. Thus, estimation of the model in (6) simplifies to multinomial logit estimation with mixed regressors (see [39], [40], [1], [41]). The presence of the dynamic terms on the right-hand side does not cause any inconsistency asymptotically. However, estimates are not asymptotically efficient unless the asymptotic variance-covariance matrix has the usual sandwich form (see [40] and [1]).

Table 3a.

Transition parameter estimates $\hat{\delta}_{jk}$ for the period 2001-2009 yearly using MLE of the dynamic MNL

Origin ↓ Destination →	EMPL	SELF	PSDRET	HOME
EMPL	7.1 (.459)	4.5 (.296)	4.8 (.607)	3.6 (.122)
SELF	5.6 (.463)	9.3 (.281)	6.1 (.590)	3.5 (.119)
PSDRET	6.2 (.536)	6.7 (.371)	11.1 (.626)	3.7 (.270)
HOME	4.1 (.495)	2.8 (.353)	3.2 (.698)	6.0 (.120)

All estimates are statistically significant at the 1% level.

Table 3b.

Transition parameter estimates $\hat{\delta}_{jk}$ for the period 2001-2009 half-yearly using MLE of the dynamic MNL

Origin ↓ Destination →	EMPL	SELF	PSDRET	HOME
EMPL	9.3 (.458)	6.0 (.288)	6.5 (.601)	4.4 (.123)
SELF	6.3 (.464)	10.8 (.280)	6.9 (.597)	4.2 (.119)
PSDRET	6.9 (.543)	7.4 (.374)	12.6 (.635)	4.5 (.277)
HOME	4.9 (.496)	3.6 (.355)	4.0 (.704)	7.5 (.121)

All estimates are statistically significant at the 1% level.

Table 3c.

Transition parameter estimates $\hat{\delta}_{jk}$ for the period 2001-2009 quarterly using MLE of the dynamic MNL

Origin ↓ Destination →	EMPLFT	EMPLPT	SELF	PSDRETHO
EMPLFT	9.0 (.179)	5.7 (.183)	7.9 (1.013)	6.2 (.336)
EMPLPT	5.2 (.200)	11.5 (.173)	9.0 (1.009)	6.7 (.332)
SELF	5.1 (.287)	8.1 (.195)	14.6 (1.009)	5.9 (.427)
PSDRETHO	5.6 (.268)	7.1 (.212)	6.8 (1.166)	12.8 (.335)

All estimates are statistically significant at the 1% level.

4.2.2 With unobserved heterogeneity (fixed effects)

This section presents estimation results of the model in (6) under the alternative hypothesis of unobserved state heterogeneity in individual preferences for work.

The existence of unobserved permanent components in individual behavior allows individuals who are otherwise homogeneous in their observed characteristics to be heterogeneous in their response probabilities $F\left(\mathbf{x}_{it}'\boldsymbol{\beta}\right)$ (see [37]). This implies that the behavior of a single individual within a group of observationally identical individuals is systematically different from group average behavior $\int F\left(\mathbf{x}'\boldsymbol{\beta}+\boldsymbol{\alpha}\right)dH\left(\boldsymbol{\alpha}|\,\mathbf{x}\right)$ marginal on the empirical distribution of heterogeneity $\boldsymbol{\alpha}$ conditional on observed characteristics $\mathbf{x}$. The presence of fixed effects in the density yields inconsistent maximum likelihood estimates unless $T\rightarrow\infty$, implying that only a limited number of observations is available to estimate the fixed effects α_i. This is known as the incidental parameters problem (see [45]). [45] have shown that this problem can be resolved if a minimum sufficient statistic τ_i existed for α_i (and did not depend on the structural parameter $\boldsymbol{\beta}$) so that the conditional density could be factorized in a way that would no longer depend on α_i :

$$f^*\left(\mathbf{y}_i|\,\boldsymbol{\beta},\tau_i\right)=\frac{f\left(\mathbf{y}_i|\,\boldsymbol{\beta},\alpha_i\right)}{g\left(\tau_i|\,\boldsymbol{\beta},\alpha_i\right)>0}$$

Moreover, [2] and [3] have shown that, under mild regularity assumptions, maximization of that conditional density would yield consistent estimates of $\boldsymbol{\beta}$.

In the bivariate case, defining $\mathbf{y}_i=\left(y_{i1},...,y_{iT}\right)$, the bivariate logit model with strictly exogenous regressors is charactered by the joint density of $\mathbf{y}_i$:

$$\Pr\left(\mathbf{y}_i|\,\alpha_i,\mathbf{x}_t;\boldsymbol{\beta}\right)=\frac{\exp\left(\alpha_i\sum_{t=1}^{T}y_{it}\right)\exp\left(\left(\sum_{t=1}^{T}y_{it}\mathbf{x}_{it}'\right)\boldsymbol{\beta}\right)}{\prod_{t=1}^{T}\left[1+\exp\left(\mathbf{x}_{it}'\boldsymbol{\beta}+\alpha_i\right)\right]} \tag{9}$$

This density is a function of the fixed effect α_i. However, since there are $\sum_t y_{it}$ outcomes equal to 1 (successes) in the T periods for individual i, [16] has shown that the number $\sum_t y_{it}$ is a minimum sufficient statistic for α_i. Define B_i the set of all possible *distinct* sequences D_{ij}

$$B_i=\left\{\,D_{ij}=\left(d_{ij1},...,d_{ijT}\right)|\,d_{ijt}=0,1\right\}$$

satistfying

$$\textstyle\sum_{t=1}^{T}d_{ijt}=\sum_{t=1}^{T}y_{it}=s$$

with

$$\begin{aligned} j &= 1,...,\frac{T!}{(T-s)!} \\ s &= 1,2,...,\tbinom{T}{s}=\frac{T!}{s!\,(T-s)!} \end{aligned}$$

The density of the $\mathbf{y}_i$ conditional on the sufficient statistics is[4]

$$\Pr\left(\mathbf{y}_{it}\middle|\sum_{t=1}^{T} y_{it} = s, \mathbf{x}_{it}; \boldsymbol{\beta}\right) = \frac{\exp\left(\left(\sum_{t=1}^{T} y_{it}\mathbf{x}_{it}'\right)\boldsymbol{\beta}\right)}{\sum_{D_{ij}\in B_i}\exp\left(\left(\sum_{t=1}^{T} d_{ijt}\mathbf{x}_{it}'\right)\boldsymbol{\beta}\right)} \tag{10}$$

The conditional maximum likelihood (CML) approach may be used to consistently estimate the 'pure' $AR(1)$ logit model with individual effects as long as there are no other regressors than the lag-dependent variable and there are at least four observations per individual including the initial time period (see [19] and [41]). Substituting $\gamma y_{i,t-1}$ for $\mathbf{x}_{it}'\boldsymbol{\beta}$ in (10) and assuming that $t = 1, ..., T$ and $T \geq 3$, the dynamic fixed effects logit model is specified as

$$\Pr\left(y_{i0} = 1|\, \alpha_i\right) = P_0\left(\alpha_i\right) \tag{11a}$$

$$\Pr\left(y_{it} = 1|\, \alpha_i, y_{i0}, ..., y_{iT}\right) = \frac{\exp\left(\gamma y_{i,t-1} + \alpha_i\right)}{1 + \exp\left(\gamma y_{i,t-1} + \alpha_i\right)} \tag{11b}$$

y_{i0} is assumed to be observed although the model is not specified in the initial period. The structural errors are independent of the initial conditions and identically distributed over time according to the logistic distribution. Assuming $T \geq 3$, the conditional log-likelihood is based on the conditional density

$$\begin{aligned} &\Pr\left(y_{i0}, ..., y_{i,T}\middle|\, y_{i0}, \sum_{t=1}^{T} y_{it}, y_{iT}\right) = \\ &= \Pr\left(y_{i1}, ..., y_{i,T-1}\middle|\, y_{i0}, \sum_{t=1}^{T} y_{it}, y_{iT}; \gamma\right) \\ &= \frac{\exp\left(\gamma \sum_{t=1}^{T} y_{it} y_{i,t-1}\right)}{\sum_{(d_{i0},...,d_{iT})\in B_i}\exp\left(\gamma\sum_{t=1}^{T} d_{it} d_{it-1}\right)} \end{aligned} \tag{12}$$

[41] derives a similar expression for the conditional density in the multivariate case. Assuming the process started at $t = 0$ and that $T \geq 3$

$$\begin{aligned} &\Pr\left(y_{i1}, ..., y_{i,T}|\, y_{i0}, n_{i1}, ..., n_{iM}, y_{iT}\right) = \\ &= \frac{\exp\sum\limits_{k>0}\sum\limits_{j}\left(\sum\limits_{t>0} 1\left\{y_{it} = k\right\} 1\left\{y_{i,t-1} = j\right\}\delta_{jk}\right)}{\sum\limits_{B}\exp\sum\limits_{k>0}\sum\limits_{j}\left(\sum\limits_{t>0} 1\left\{y_{it} = k\right\} 1\left\{y_{i,t-1} = j\right\}\delta_{jk}\right)} \end{aligned} \tag{13}$$

where

$$n_{ik} = \sum_{t=1}^{T} 1\left\{y_{it} = k\right\}$$

[4] Since sequences with $s = 0$ or T make no contribution to the likelihood function, only $T - 1$ alternative sets are appropriate for conditioning.

and

$$B = \left\{ b = (y_{i0}, ..., y_{i,T}) | \forall k > 0; \sum_{t=1}^{T} 1\{y_{it} = k\} = n_{ik} \right\}$$

This model is identified up to $M^2 - (2M - 1)$ transition parameters δ_{jk} since the normalization constraint requires that the last row and column of the transition matrix corresponding to the reference state were set to zero.

Tables 4a-c show CML estimates of the model in (13) implied by (6).

Table 4a.

Transition parameter estimates $\hat{\delta}_{jk}$ for the period 2001-2009 yearly conditional on unobserved individual heterogeneity

Origin ↓ Destination →	EMPL	SELF	PSDRET	HOME
EMPL	9.3 (3.106)	5.1 (.848)	5.3 (2.504)	3.4 (.410)
SELF	6.6 (3.026)	9.0 (.750)	5.6 (2.309)	2.7 (.417)
PSDRET	6.2 (3.002)	5.2 (1.419)	8.0 (2.508)	*1.8* (1.134)
HOME	*5.5* (3.088)	2.8 (.855)	*2.0* (2.260)	5.1 (.425)

$CLE: N = 2026$ and $\ln L = -0.2474$.

Entries $(3, 4)$ and $(4, 1)$ are statistically significant at the 11% and 7% level, respectively. Entry $(4, 3)$ is statistically insignificant.

All other estimates are statistically significant at $\leq 4\%$ level.

Table 4b.

Transition parameter estimates $\hat{\delta}_{jk}$ for the period 2001-2009 half-yearly conditional on unobserved individual heterogeneity

Origin ↓ Destination →	EMPL	SELF	PSDRET	HOME
EMPL	10.6 (.874)	6.2 (.574)	6.2 (1.442)	4.0 (.205)
SELF	7.0 (0.928)	10.2 (.480)	6.5 (1.478)	3.3 (.235)
PSDRET	*5.9* (1.135)	5.4 (.920)	9.1 (1.506)	2.5 (.752)
HOME	*4.1* (.924)	2.3 (.661)	*2.1* (1.536)	6.4 (.225)

$CLE: N = 2097$ and $\ln L = -0.3764$.

Entries $(3, 1), (4, 1)$ and $(4, 3)$ are statistically significant at the $1\%, 4\%$ and 3% level, respectively.

All other estimates have a p-value close to zero.

Table 4c.
Transition parameter estimates $\hat{\delta}_{jk}$ for the period 2001-2009 quarterly conditional on unobserved individual heterogeneity

Origin ↓ Destination →	EMPLFT	EMPLPT	SELF	PSDRETHO
EMPLFT	10.0 (.376)	6.4 (.460)	9.5 (1.244)	6.9 (.708)
EMPLPT	6.0 (.499)	11.2 (.555)	10.0 (1.152)	7.2 (.658)
SELF	7.5 (.597)	10.9 (.523)	21.2 (1.560)	9.4 (.890)
PSDRETHO	6.0 (.611)	7.1 (.578)	*7.3* (3.217)	11.6 (.682)

$CLE : N = 2261$ and $\ln L = -.5209$.
Entry $(4, 3)$ is statistically significant at the 2% level.
All other estimates have a p-value close to zero.

The conditional logit estimates in *Tables 4a-c* suggest that the higher the frequency, or the shorter the time-interval, the higher the degree of state dependence. Moreover, irrespective of the frequency, state dependence, as measured by the transition parameters, is significant in all dimensions. Since all transition parameter estimates are positive, the odds of being in any employment state with respect to non-participation whatever the lagged employment state are larger than when the lagged state is non-participation. In other words, non-participation is less likely to be either the origin or the destination state for any employment state. Moreover, the delay to enter other employment states after a non-participation spell is significantly larger than staying active even when fixed characteristics, such as preferences for work, have been accounted for.

4.2.3 Structural tests

The estimation results of *Tables 3a-c* and *4a-c* may be contrasted by means of the Hausman statistic testing the null hypothesis of no unobserved state heterogeneity in the employment choices of the 50-64 sample respondents. A rejection of the model under the null would imply true state dependence in choices conditional on individual preferences for work.

In general, the hypotheses may be stated as

$$\begin{aligned} H_0 &: \alpha_i = \alpha \rightarrow \delta^i_{jk} = \delta_{jk} \ \forall i \\ H_a &: \alpha_i \neq \alpha \rightarrow \delta^i_{jk} \neq \delta_{jk} \ \forall i \end{aligned}$$

This amounts to testing the homogeneity assumption of the underlying structural model. It is not possible to test the null hypothesis using the likelihood ratio test because the two likelihoods are not comparable. Under the null, both estimators are consistent but the unconditional one is asymptotically efficient as well. Under the alternative, the unconditional estimator is inconsistent whereas the conditional one is both consistent and asymptotically efficient (see

[31]). The Hausman statistic is distributed as a χ^2 with k degrees of freedom. The variance-covariance difference should be positive definite otherwise the value of the statistic is zero. The general formula of the statistic is

$$\chi^2(k) = \left(\hat{\boldsymbol{\delta}}_{CML} - \hat{\boldsymbol{\delta}}_{ML}\right)' \left(\widehat{AV}_{CML} - \widehat{AV}_{ML}\right)^{-1} \left(\hat{\boldsymbol{\delta}}_{CML} - \hat{\boldsymbol{\delta}}_{ML}\right)$$

Leaving non-participation out of the picture, state and duration dependence may be diagnosed by performing Hausman tests for certain dimensions (rows, columns) of the estimated transition matrix (see [41], [48], [4]). If the coefficients in a row are equal, the corresponding origin state would not affect the access rate to any employment state. Similarly, if the coefficients in a row are equal, the corresponding destination state would be equally accessible from every other employment state.

Tables 5a-c show the test statistics for the estimated transition matrix as whole as well as for its rows (holding the origin state constant).

Table 5a.

Yearly Hausman statistics at the 5% significance level

Overall Hausman statistic: $34.60 > \chi^2_c(16) = 26.3$

Origin state ↓	Hausman statistic	Critical value
EMPL	*7.73*	$\chi^2_c(4) = 9.49$
SELF	101.92	$\chi^2_c(4) = 9.49$
PSDRET	19.54	$\chi^2_c(4) = 9.49$
HOME	10.09	$\chi^2_c(4) = 9.49$

Variance-covariance difference is PD.

Table 5b.

Half-yearly Hausman statistics at the 5% significance level

Overall Hausman statistic: $66.54 > \chi^2_c(16) = 26.3$

Origin state ↓	Hausman statistic	Critical value
EMPL	17.76	$\chi^2_c(4) = 9.49$
SELF	132.61	$\chi^2_c(4) = 9.49$
PSDRET	34.35	$\chi^2_c(4) = 9.49$
HOME	19.83	$\chi^2_c(4) = 9.49$

Variance-covariance difference is PD.

Table 5c.

Quarterly Hausman statistics at the 5% significance level

Overall Hausman statistic: $131.59 > \chi^2_c(16) = 26.3$

Origin state ↓	Hausman statistic	Critical value
EMPLFT	54.07	$\chi^2_c(4) = 9.49$
EMPLPT	86.82	$\chi^2_c(4) = 9.49$
SELF	78.12	$\chi^2_c(4) = 9.49$
PSDRETHO	59.24	$\chi^2_c(4) = 9.49$

Variance-covariance difference is PD.

Both overall and individual row test values exceed the corresponding critical values in all but one cases. Overall significance suggests that individual-specific unobserved heterogeneity should be included in the model validating the dynamic fixed-effects reduced form as the appropriate model. Row heterogeneity is evidence in favor of state and duration dependence in employment choices as it would imply barriers to entry in the labor market contradicting the second structural assumption. Indeed, observing *Tables 4a-c*, being in one employment state decreases the probability of being in a different one the following period, as indicated by the relatively lower estimates as we move away from the diagonal element of each row of the estimated transition matrix. This indicates that employment spells (full-time, part-time) have relatively longer duration. For instance, $10.6 > 6.2 = 6.2 > 4.0$ in row 1 $\{10.6, 6.2, 6.2, 4.0\}$ of *Table 4b*.

5 Long-run determinants

5.1 Quasi-random effects estimation

(Work-in-progress).

6 Conclusion

This paper has attempted to establish the degree to which the employment transitions of older persons, namely of the 50-64 age cohort, exhibit true state and duration dependence conditional on individual preference heterogeneity as a function of the current labor market state. Estimation results provide evidence in favor of true state dependence implying that preferences for work of the 50-64 ELSA respondents are indeed heterogeneous. Moreover, the longer the duration of an inactivity spell (non-participation), the greater the delay to reenter the labor market. Holding non-participation constant, employment spells have relatively longer duration. It remains to determine the long-run determinants of this preference heterogeneity *(work-in-progress)...*

References

[1] Amemiya, T. (1985). *Advanced Econometrics.* Cambridge: Harvard University Press.

[2] Andersen, E.B. (1970): 'Asymptotic Properties of Conditional Maximum Likelihood Estimators,' *Journal of the Royal Statistical Society, Series B*, **32**, 283-301.

[3] Andersen, E.B. (1973). *Conditional Inference and Models for Measuring.* Köbenhavn: Mentalhygiejnish Farlag.

[4] Anderson, T.W. and L.A. Goodman (1957): 'Statistical Inference about Markov Chains,' *Annals of Mathematical Statistics*, **28**, 89-110.

[5] Behncke, S. (2009): 'How Does Retirement Affect Health?,' IZA DP No. 4253.

[6] Belloni, M. and R. Alessie (2008): 'The importance of financial incentives on retirement choices: New evidence for Italy,' *Labour Economics*, **16**, 578–588.

[7] Berkovec, J. and S. Stern (1991): 'Job Exit Behavior of Older Men,' *Econometrica*, **59**, 189-210.

[8] Blau, D.M. and R.T. Riphahn (1999): 'Labor Force Transitions of Older Married Couples in Germany,' *Labor Economics*, **6**, 5-46.

[9] Bloemen, H.G. (2008): 'Private Wealth and Job Exit at Older Age: A Random Effects Model', IZA Discussion Paper No. 3386.

[10] Blundell, R.W., C. Meghir, and S. Smith (2002): 'Pension Incentives and the Pattern of Early Retirement,' *The Economic Journal*, **112**, C153-C170.

[11] Bockerman, P. and P. Ilmakunnas (2009): 'Unemployment and Self-assessed Health: Evidence from Panel Data,' *Health Economics*, **18**, 161-179.

[12] Boheim, R. and M. Taylor (2000): 'Unemployment duration and exit rates in Britain,' ISER Working Paper No. 50.

[13] Börsch-Supan, A., A. Brugiavini and E. Croda (2009): 'The role of institutions and health in European patterns of work and retirement,' *Journal of European Social Policy*, **19**, 341-358.

[14] Bound, J., M. Schoenbaum, T.M. Stinebrickner and M. Waidmann (1999): 'The Dynamic Effects of Health on the Labor Force Transitions of Older Workers,' *Labour Economics*, **6**, 179–202.

[15] Cai, L (2010): 'The Relationship between Health and Labour Force Participation: Evidence from a Panel Data Simultaneous Equation Model,' *Labour Economics*, **17**, 77–90.

[16] Chamberlain, G. (1980): 'Analysis of Covariance with Qualitative Data,' *Review of Economic Studies*, **47**, 225-238.

[17] Chamberlain, G. (1982): 'Multivariate Regression Models for Panel Data,' *Journal of Econometrics*, **18**, 5-46.

[18] Chamberlain, G. (1984): 'Panel Data,' in Z. Griliches and M. Intriligator (Eds.) *Handbook of Econometrics II*, 1247-1318. Amsterdam: North-Holland.

[19] Chamberlain, G. (1985): 'Heterogeneity, Omitted Variable Bias and Duration Dependence,' in *Longitudinal Analysis of Labor Market Data*, 3-38, edited by J.J. Heckman and B. Singer. Cambridge, UK: Cambridge University Press.

[20] Chamberlain, G. (1993): 'Feedback in Panel Data Models,' *mimeo*. Department of Economics, Harvard University.

[21] Chintagunta, P., E. Kyriazidou and J. Perktold (2001): 'Panel Data Analysis of Household Brand Choices,' *Journal of Econometrics*, **103**, 111-153.

[22] Contoyannis, P. and N. Rice (2001): 'The Impact of Health on Wages: Evidence from the British Household Panel Survey,' *Empirical Economics*, **26**, 599–622.

[23] Currie, J. and B.C. Madrian (1999): 'Health, Health Insurance and the Labour Market,' in *Handbook of Labour Economics III*, 3309–3416, edited by O. Ashenfelter and D. Card. Amsterdam: North-Holland.

[24] Disney, R., C. Emmerson and M. Wakefield (2006): 'Ill health and retirement in Britain: A panel data-based analysis,' *Journal of Health Economics*, **25**, 621-649.

[25] Duval, R. (2003): 'The Retirement Effects of Old-Age Pension and Early Retirement Schemes in OECD Countries,' OECD Economics Department Working Papers, No. 370, OECD publishing, OECD.

[26] Fougère, D. and T. Kamionka (1992): 'Un modèle markovien du marché de travail,' *Annales d'Économie et de Statistiques*, **27**, 149-188.

[27] Frederiksen, A., Bo E. Honoré and L. Hu (2007): 'Discrete Time Duration Models with Group-Level Heterogeneity,' *Journal of Econometrics*, **141**, 1014-1043.

[28] García-Gómez, P. and A. López-Nicolás (2006): 'Health Shocks, Employment and Income in the Spanish Labour Market,' *Health Economics*, **15**, 997–1009.

[29] García-Gómez, P., A.M. Jones, and N. Rice (2010): 'Health Effects on Labour Market Exits and Entries,' *Labour Economics*, **17**, 62-76.

[30] Gruber, J. and D.A. Wise (Eds.) (1999, 2004, 2007). *Social Security Programs and Retirement around the World*. Chicago: University of Chicago Press.

[31] Greene, W.H. (2003). *Econometric Analysis* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.

[32] Heckman, J.J. (1978): 'Simple Statistical Models for Discrete Panel Data Developed and Applied to Test the Hypothesis of True State Dependence against the Hypothesis of Spurious State Dependence,' *Annales de l'INSEE*, **30**, 227-269.

[33] Heckman, J.J. (1981): 'Heterogeneity and State Dependence,' in S. Rosen (Ed.) *Studies in Labor Markets*, 91-139. University of Chicago Press.

[34] Heckman, J.J. (1981): 'The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process,' in *Structural Analysis of Discrete Data with Econometric Applications*, edited by C.F. Manski and D. McFadden, 179-195 (Chapter 4). Cambridge, MA: MIT Press.

[35] Honoré, Bo E. and E. Kyriazidou (2000): 'Panel Data Discrete Choice Models with Lagged Dependent Variables,' *Econometrica*, **68**, 839-874.

[36] Honoré, Bo E. and E. Tamer (2006): 'Bounds on Parameters in Panel Dynamic Discrete Choice Models,' *Econometrica*, **74**, 611-629.

[37] Hsiao, C. (2003). *Analysis of Panel Data* (2nd ed.). Cambridge: Cambridge University Press.

[38] Lazear, E. (1986): 'Retirement from the labor force,' in *Handbook of Labor Economics I.*, edited by O. Ashenfelter and R. Layard, 305–355. Amsterdam: Elsevier Science Publishers BV.

[39] McFadden, D. (1973): 'Conditional Logit Analysis of Qualitative Choice Behaviour,' in *Frontiers in Econometrics*, edited by P. Zarembka, 105-142. New York: Academic Press.

[40] Maddala, G.S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics.* ESM 3. Cambridge: Cambridge University Press.

[41] Magnac, T. (2000): 'Subsidised Training and Youth Employment: Distinguishing Unobserved Heterogeneity from State Dependence in Labour Market Histories,' *The Economic Journal*, **110**, 805-837.

[42] Maurer, J., R. Klein and F. Vella (2007): 'Subjective Health Assessments and Active Labour Market Participation of Older Men: Evidence from a Semiparametric Binary Choice Model with Nonadditive Correlated Individual-Specific Effects,' IZA DP No. 3257.

[43] Marmot, M., J. Banks, R. Blundell, C. Lessof and J. Nazroo (2004). ELSA Report. Institute for Fiscal Studies.

[44] Mortensen, D. and K. Neuman (1984): 'Choice or Chance? A Structural Interpretation of Individual Labour Market Histories,' in *Studies in Labour Market Dynamics*, edited by K. Neuman and N. Westergaard-Nielsen, 98-131. Heidelberg: Springer-Verlag.

[45] Neyman, J. and E.L. Scott (1948): 'Consistent Estimates Based on Partially Consistent Observations,' *Econometrica*, **16**, 1-32.

[46] Rice, N., J. Roberts and A. Jones (2006): 'Sick of work or too sick to work? Evidence on health shocks and early retirement from the BHPS,' Health, Econometrics and Data Group Working Paper 2006/13. University of York.

[47] Riphahn, R. (1999): 'Income and Employment Effects of Health Shocks. A Test Case for the German Welfare State,' *Journal of Population Economics*, **12**, 363–389.

[48] Ruud, P. (1984): 'Tests of Specification in Econometrics,' *Econometric Reviews*, **3**, 211-242.

[49] Smith, J.P. (2004): 'Unravelling the SES Health Connection,' The Institute for Fiscal Studies, Working Paper WP04/02, London.

[50] Stewart, J.M. (2001): 'The Impact of Health Status on the Duration of Unemployment Spells and the Implications for Studies of the Impact of Unemployment on Health Status,' *Journal of Health Economics*, **20**, 781–796.

[51] Stock, J.H. and D.A. Wise (1990): 'Pensions, the option value of work, and retirement,' *Econometrica*, **58**, 1151-80.