Bolano, D. & M. Haynes (2016)

Methodological approaches to profiling and modelling disadvantaged employment pathways. An application to employment trajectories in Australia

in G. Ritschard & M. Studer (eds), Proceedings of the International Conference on Sequence Analysis and Related Methods, Lausanne, June 8-10, 2016, pp 851-862.

Methodological approaches to profiling and modelling disadvantaged employment pathways. An application to employment trajectories in Australia

Danilo Bolano and Michele Haynes

Abstract This paper investigates the employment pathways of men and women in Australia and profiles the characteristics of individuals who are at risk of disadvantage defined by unstable employment histories and frequent transitions into unemployment. The paper focuses on the transitions of respondents, aged 15-64 years, between different employment states over a span of 13 years (2001-2013) using panel data from the Household, Income and Labour Dynamics in Australia survey. To describe employment trajectories and to analyse the likelihood of transition from one employment state to another, we will rely on sequence analysis and two probabilistic models that might account for state dependence: dynamic multinomial logit random effects models and Markov models. Sequence analysis was used to identify typology of employment pathways and the associated sociodemographic characteristics and intergenerational links. The preliminary results confirm the presence of gender differences in employment. Women are more likely to be employed part time or not in the labour force while men are more likely to experience stable employment trajectories of full time work. Women also experience a slightly higher proportion of employment transitions. Among the sociodemographic and background factors considered, mature age workers, those with health problems and less educated parents, in particular father-son and mother-daughter links, are mainly at risk of experiencing unstable employment pathways.

1 Introduction

Trends in unemployment are of interest due to the financial and social costs borne by individuals. Some of the cost are absorbed by welfare payments but many of which

Danilo Bolano
Institute for Social Science Research, The University of Queensland, e-mail: d.bolano@uq.edu.au

Michele Haynes
Institute for Social Science Research, The University of Queensland e-mail: m.haynes@uq.edu.au

are faced directly by families and communities who may not have the capacity to absorb them. Unemployment not only means a loss of economic security but also feeds poverty and social dislocation as the ties of civil society are severed. In general, unemployment removes an important set of social relationships without which many people have little support or security in dealing with often profound economic hardship. The longer the duration in transition from unemployment to employment, the more these hardships are exacerbated. Therefore unstable employment histories with long periods of unemployment can be considered as a measure of social disadvantage.

Unstable employment history can lead to labor market disadvantages too. The longer the period of time spent in unemployment, the lower an individual's likelihood of gaining employment. On the supply-side of the labor market, employers may be reluctant to employ the long-term unemployed, in which case economic policies aimed at reducing unemployment may have limited success (ABS, 2001). And on the supply side, absences from employment are associated with loss of human capital, under use and underdevelopment of skills (Vickery, 1999), limited occupational advancement (Burbidge, 2005). These can lead to the deterioration of labour market prospects, or scarring, in the form of relatively poor prospects of finding work (ABS, 2001; Saunders and Brown, 2004), loss of income whilst unemployed and lower wages on return to the workforce (Arulampalam, 2001; Green and Leeves, 2009; Stevens, 1997).

A range of personal characteristics are known to be associated with transition in and out the labor force and the duration of unemployment spells. In particular, we focus on the difference in employment trends for men and women. Women have often indicated a preference for part-time work as a way of attempting to combine child bearing and child rearing, home and non-child related caring responsibilities, for which women still bear the overwhelming responsibility (ABS, 2009; Hakim, 2000).

The paper has three objectives. First, to describe the different employment pathways of men and women in Australia. Second, to profile individuals at greater risk of disadvantage by demographic characteristics and family background, where disadvantage is defined in terms of unstable labor market history and transitions out of the labor force. Third, to review and compare statistical approaches to analyze employment trajectories. These approaches include sequence analysis, generalized linear mixed model in the form of a dynamic multinomial logit random effects model with lagged dependent variable to account for state dependence and Markov-based models.

2 Data and Methods

We use data from the first thirteen waves of the Household, Income and Labour Dynamics in Australia (HILDA) panel survey. It is a household-based panel study which began in 2001. Information are collected annually from each person in the household aged 15 and more. The first wave was carried out from a sample of 13,969 individuals from 7,682 households (Watson and Wooden 2002).

This paper investigates employment trajectories and potential sociodemographic and family determinants of disadvantaged employment transition pathways. Therefore, we select only individuals of working age (15 to 65 years old) who participated in at least to two waves of the HILDA survey since Wave 1 (mean number of observations per individual = 9.92). The final sample consists of 10,941 respondents (52.14% of women, n=5,705). The mean age at the first observation was 38.5 years (median = 39).

The response variable is employment status defined by five categories: 'Being Full Time Student' (EduFT - defined as those who are enrolled in full time education and not employed), 'Employed Full Time' (FT), 'Employed Part Time' (PT), 'Looking for a job (LookingJob)', 'Not in the Labor Force' (NotLabourForce). Non responses have been included as an additional category ('Non Respond'). In the final version of the paper, we will analyse in more details the missingness patterns.

32.8% of respondents (3,584 individuals) are in the same employment state over time and 67.5% of them (2,419 individuals) declare to be employed full time at each wave (it is worth recalling now that the number of observations for each individual can vary from 2 up to 13). Pooling together the 13 waves of HILDA survey (Table1), we observe the predominance of situations of full time employment (51.2%) and part time jobs (21.7%). However, the share of times spent outside the labor force is also reasonably high (20.27%).

Table 1 Distribution of states, pooled waves 1-13

	Frequency	Percentage
Full time student (EduFT)	1,776	1.64
Full time employed (FT)	55,557	51.19
Part time employed (PT)	23,503	21.66
Looking for a job (LookingJob)	2,953	2.72
Not in labour force (NotLabourForce)	21,994	20.27
Non respond (NonRespond)	2,740	72.53
Total data points	108.523	

To investigate potential sources of disadvantage in employment transitions, a set of demographic and human capital characteristics of the respondents, health condition, socio-economic indicators (Socio-Economic Index for Areas and AUSEI), relevant life transitions (transition to parenthood) and socio-economic backgrounds (parental educational level, socio-economic status, parents occupation) will be included in the final analysis. The paper aims not only to identify the key factors associated to employment transitions between working statuses but also to review and compare different approaches for analyzing individual patterns. We will consider both descriptive tools (sequence analysis) as in this abstract and probabilistic models as Markov models and generalized linear mixed models with random effects (see e.g., Haynes et al. 2008).

3 Clustering using Sequence Analysis and Preliminary Results

With a holistic approach (Billari, 2001), the unit of interest is the entire individual trajectory and holistic studies mainly rely on sequence analysis (SA, Abbott, 1995). Using a sequence analysis approach, working trajectories are described as an ordered sequence of labour market states. Looking at the entire set of individual transitions among states occurred over time (potentially over an entire lifetime career) as a unique sequence of observations, sequence analysis can discover, describe and explain different individual life course patterns being more informative then focusing on one single working transition.

SA has been applied to study different types of social processes. Studies on work trajectories and pathways from school to work include McVicar and Anyadike-Danes, 2002; Martin et al., 2008, Fasang 2010. As far as we know, the only extensive study on employment trajectories in Australia using sequence analysis is Fry and Boulton (2013). This paper extends their work in several ways. The paper uses all the waves of HILDA survey available at the time of writing this abstract (13 waves, 2001-2013), compares different competitive statistical models and takes a life course approach. The papers controls for demographic characteristics, life events and socio-demographic background to identify the factors that are more likely associated to employment mobility and situation of working disadvantages.

The length of employment history considered for each individual can vary from 2 up to 13 years. For 5,246 individuals—47.9% of our sample—we have complete sequences. The most frequent pattern is being full time employed. Such stable working condition concerns 1,253 individuals. Only 220 respondents (2% of the entire sample) are outside the labour force and not in education during the entire period from 2001 and 2013.

Table 2 Most common working subsequences. Ordered by frequency for the whole sample.

Subsequence	Entire sample (n=10,941)		Female (n=5,705)		Male (n=5,236)	
	Freq	Percent	Freq	Percent	Freq	Percent
Full Time Employed over the 13 waves	1253	11.45	260	4.56	993	18.96
FT- FT	226	2.07	73	1.28	153	2.92
Not in the labour force over the 13 waves	220	2.01	169	2.96	51	0.97
NotLabourForce NotLabourForce	187	1.71	122	2.14	65	1.24
FT FT - FT	153	1.4	50	0.88	103	1.97
NotLabourForce- NotLabourForce- NotLabourForce	128	1.17	79	1.38	49	0.94
Part Time Employed over the 13 waves	96	0.88	96	1.68	-	-

The stability in working trajectories in Australia is confirmed looking at the transitions between labour market states over two consecutive waves (Figure 1) and the most common (sub)sequences reported in Table 2. On average, people experiences less than 2 working transitions (mean of 1.848) over the life span considered and in 3,952 cases (45.19% of the total transitions) the respondent remains in the same labour market states (i.e., we observe no transitions) over two consecutive periods.

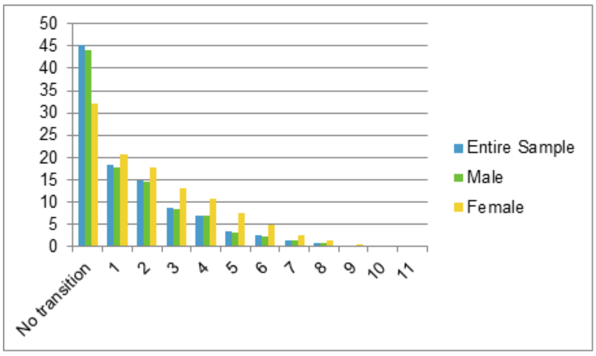

Fig. 1 Distribution of working transitions over two consecutive periods. Entire sample and divided by gender.

It is not surprising that women seem to follow more complex working trajectories experiencing a higher percentage of working transitions (see Figure 1) and spending more time outside the labour market or in a part time job (Figure 2). As reported in Table 2, only 4.56% of women declare to be full time employed from 2001 and 2013 against 19% of men. And, more than 6% of women spent at least two consecutive periods outside the labour force. In particular, almost 3% of women between 15 and 65 years old included in our sample never entered in the labour market in any of the waves considered.

Figure 3 shows graphically the working trajectories of 10 individuals (top right panel for the entire sample and bottom panels by gender) and the cross sectional distribution of states for the whole sample (top left panel). Figure 4 shows instead the differences between genders in the most common working trajectories. The R package TraMineR (Gabadinho et al., 2011) has been used for the analysis. As discussed before, respondents tend to be in the same state over time but men are more frequently in a stable full time position (violet in the graph) while women are commonly either outside the labour force (blue state) or in a part time job over time (red state). Moreover, the pictures seem suggesting a greater between-individual variations among employment trajectories for women.

Sequence analysis can be used not only to describe and represent trajectories but also to compare patterns and cluster individuals in homogeneous groups according to their pathways. Using the Optimal Matching distance as a measure of similarity between working trajectories (Abbott and Forrest 1986), we have identified four typologies for each gender. For women, the first cluster represents those are most time in a part time job (1,406 individuals). The second cluster, that is the most common one (n=1,913), groups women who are frequently outside the labor force during the period of observations. Women employed in full time o jobs are represented in

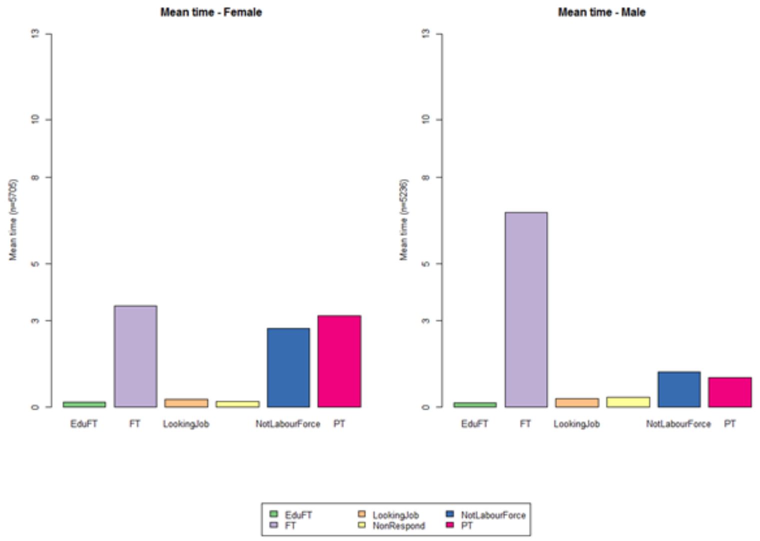

Fig. 2 Mean time (number of waves) spent in each working status. Divided by gender.

group 3 (n=1,664). The remaining 722 female participants characterized by higher variability in working trajectories are grouped in cluster 4.

For men, the clusters are less differentiate with cluster two (n=1,292) and three (n=1,729) representing situations of long term full time employment. Cluster 4 includes 578 individuals who are outside labor force for several years. Finally, cluster 1 groups shorter sequences and individual who has several periods of part time job. Further analysis will be performed accounting for the patterns of missing responses.

More sophisticated techniques in sequence analysis allow to analyze the factors that better discriminate between different pathways. Using discrepancy analysis, we might test for instance the share of discrepancy between employment trajectories explained by human capital (see Studer et al. 2011 for a detailed discussions on discrepancy analysis). According to our first results (not shown in this abstract), the education level of respondents explains 1.3% of discrepancy in working trajectories in Australia among male respondents and 1.6% among women. Similarly, the socio-demographic background explain less than 2% of discrepancy between pathways. On the other hand, the age of respondents explains 10% of the differences among trajectories observed for men.

To identify profiles at risk of having unstable and uncommon working trajectories a regression tree model can be used. The regression trees (Figure 6 and Figure 7) allow to test the association between the observed working trajectories and a set of individual characteristics. The socio-demographic background plays a central role in explaining the differences observed in the trajectory (i.e., between individual variations) determining the first split in the tree. Individuals with a better social background (i.e., a parent with a minimum level of education) experiences longer

Fig. 3 Graphical representations of working trajectories in Australia. Cross-sectional distribution (top panel left-hand side). 10 random working trajectories (top right). Individual sequences by gender on the bottom.

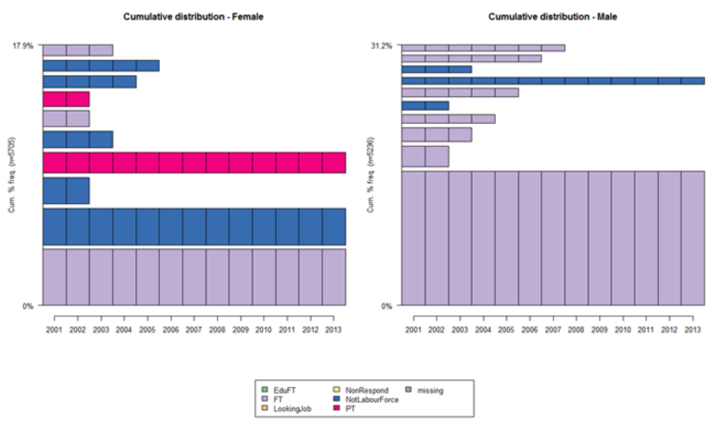

Fig. 4 Most common working trajectories. 2001-2013. Divided by gender.

periods of full time or part time job (Figure 6). The age of the respondents is, as expected, another key factor. Older cohorts are more likely retired or outside labor force (the blue state in the graphs) with respect to respondents between 14 and 54

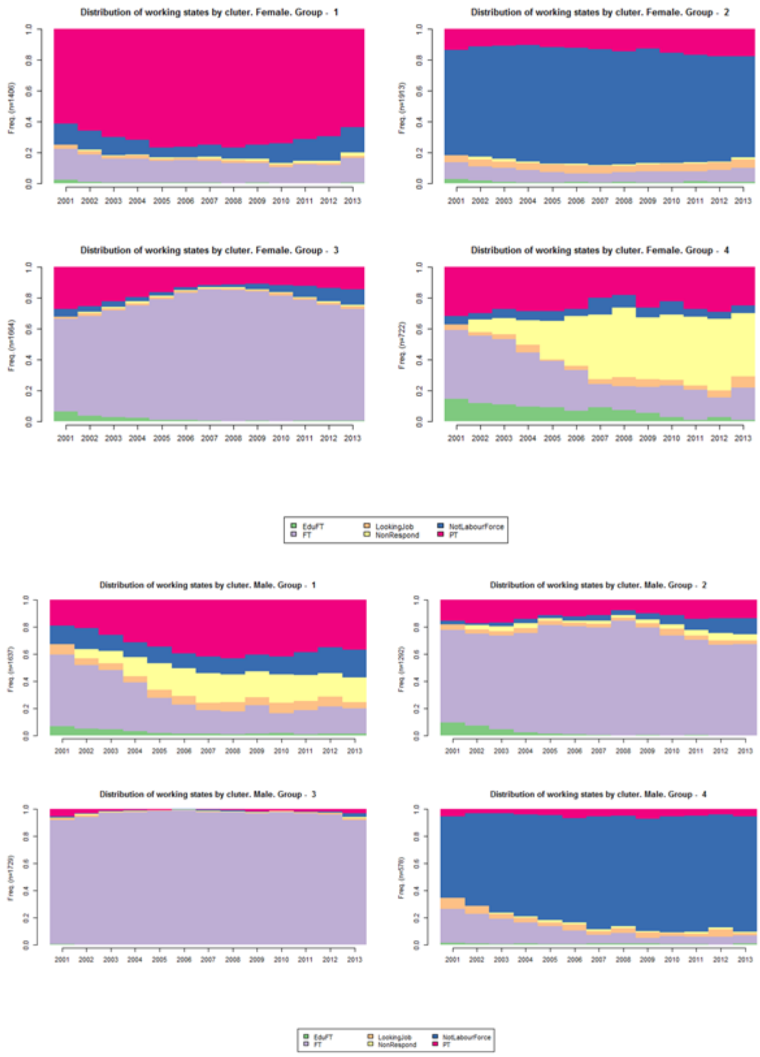

Fig. 5 Cross sectional distribution of working states by cluster. By gender.

years old. Finally, the health condition has a strong influence on labor force participation. For instance, looking at the bottom right corner of Figure 6, a male respondent aged 14 - 54 with a father with a minimum level of education, is more likely outside the labor force or in part time job if reports a fair or poor health condition.

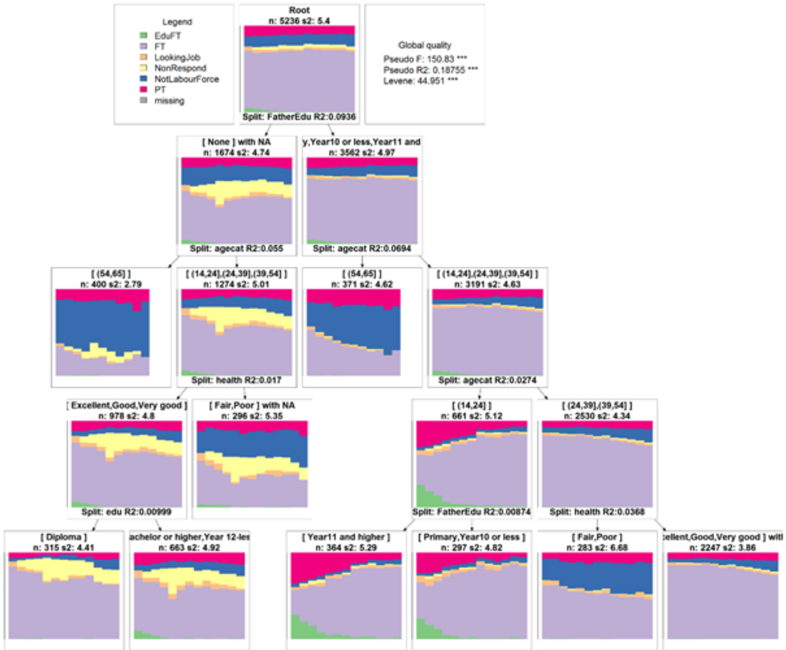

Fig. 6 Regression tree for working trajectories. Male respondents (n=5,236)

Thus these preliminary results confirm that relevant gender differences in labor force participation and stability in working trajectories still exist nowadays in Australia. It is interesting to notice another gender-related factor with respect to the family background of the respondents. The results show that a significant intergenerational relationship through same gender parent-child. For a male respondent, it is the level of education of the father that is the more associated with a stable working trajectory. For a female respondent, the participation to the labor force seems associated with the level of education of the mother. Such father-son and mother-daughter relationship should be investigated further also taking into account the missing information on parent level of education. On the other hand, educational level seems to be less relevant such as area-level indicators of disadvantages (SEIFA). However more sophisticated analysis will be performed in the final version of the paper to explore better on the effect of sociodemographic and background characteristic on working trajectories in Australia.

As briefly shown, Sequence Analysis allows to represent, describe and identify the factors associated to unusual and unstable employment trajectories. However, as most data mining techniques, it does not make any assumption about the 'social process' that have generated the trajectories. We will then explore alternative (stochastic) methods such as Markov models (see Billard 2001 for an introduction) and generalized linear mixed effect models (GLMM. See for instance Wooldridge 2002, Haynes et al. 2008). Using SA and stochastic models we might be able to

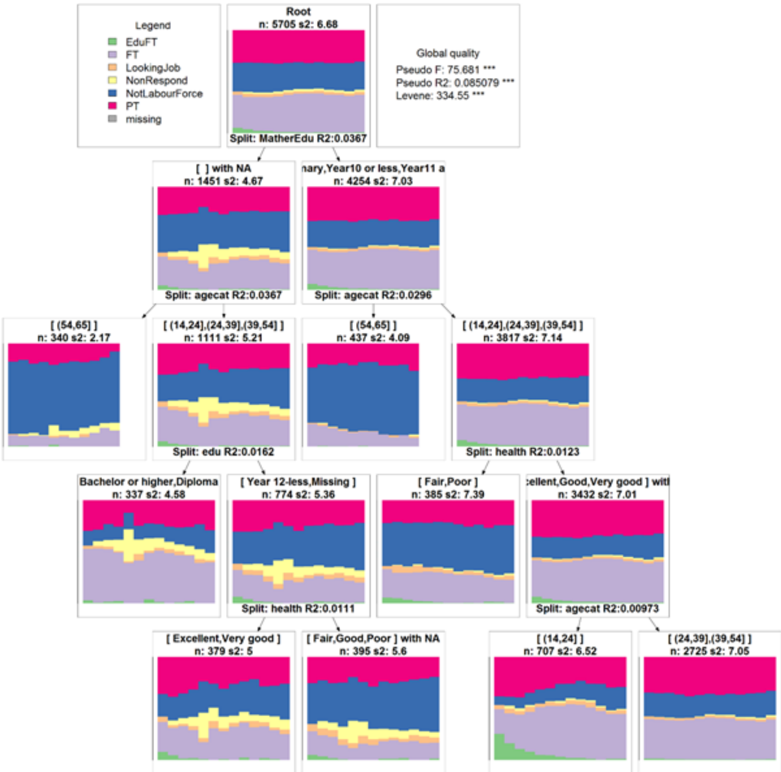

Fig. 7 Regression tree for working trajectories. Female respondents (n=5,705)

identify both between-individual differences in employment trajectories and within individual variations.

In the Markovian perspective for example, life trajectories are considered as the result of a stochastic process in which the probability of occurrence of a particular state or event depends on the sequence of states observed so far and, eventually, on a set of covariates. In other words, considering life trajectories as sequences of mutually exclusive states—e.g., sequences of employment statuses—the Markovian perspective focuses on the successive transitions and attempts to depict the life history of an individual by means of the probabilities to switch to the different states of interest given the state history lived so far. The focus is more on the transitions than on the entire trajectories. By the means of probabilistic models, we might then investigate the effect of the socio-economic background of the respondents in transitions out of unemployment (i.e., the probability of observing the transition 'Out of labor force - Employed') and the accumulation of human capital as potential protective factor against 'falling out' the labor force (i.e., of observing the transition 'Employed - Not Employed'). Then, we might profile individuals at greater risks of disadvantages.

The generalized linear mixed model (GLMM) often used to analyse nominal variables with repeated observations is the dynamic multinomial logit random effects model. For such model, the response distribution is defined conditionally on the random effects that are assumed to arise from a multivariate distribution. For a conventional multinomial logit model, let Y_{it} denote the tth observation for individual i with J possible states (6 working states in our case), $Pr(Y_{it} = j|X_{it})$ is the probability of being in state j at time t given a set of exploratory variables. Among them, we will also include the employment status in the previous wave (e.g., lagged dependent variable) in order to include explicitly the time dependence between observations.

The multinomial model can be expressed as follows

$$\pi_{itj} = Pr(Y_{it} = j|X_{it}) = \frac{e^{X_{it}\beta j}}{\sum_{k=1}^{J} e^{X_{it}\beta k}}$$

The logit model pairs each response category with an arbitrary baseline category. For instance if the first response (state j_1) is set as reference, the multinomial logit model will have the form

$$\log\left(\frac{\pi_{itj}}{\pi_{it1}}\right) = X_{it}\beta j \quad j = 2,3,...,J$$

We might interpret the logit as sort of utility function in a decision making process. We define the utility of choosing a particular response, a certain employment state, by the random variables U_{itj} $(j = 1,,J)$, with the function $U_{itj} = X_{it}\beta_j + e_{itj}$. An individual will choose the response j, for example he/she will decide to not participate to the labor market, if the utility in staying in this state is the greatest. In other words if $U_{itj} = \max_{1\leq k\leq J} U_{itk}$.

We will also consider another model specification introducing individual-specific random effects to model sprurious dependence. The random effects $\alpha_i = \{\alpha_{ij}, ..., \alpha_{iJ}\}$ capture non-observable individual effects that are specified to arise from a multivariate normal distribution. The multinomial logit random effect model is then defined as

$$\log\left(\frac{\pi_{itj}}{\pi_{it1}}\right) = X_{it}\beta j + Z_{ij}\alpha_{ij} \quad j = 2,3,...,J$$

Where Z_{ij} denote a vector of coefficients for the random effects.

The paper intends not only to illustrate different methodological approaches for analyzing longitudinal data but also to show how such approaches, investigating the socio-demographic and background factors leading to situation of disadvantages, may provide guides to policy makers who are interested in addressing labor market disadvantages.

References

1. Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology* 21, 93-113.
2. Abbott, A. and Forrest (1986). Optimal Matching for Historical Sequences. *Journal of Interdisciplinary History*. 16(3). 471-794.
3. ABS. (2001). Duration of Unemployment: Recent Definitional Changes. *Canberra: Australian Bureau of Statistics. Cat. No. 6203.0*.
4. ABS. (2009). Trends in Household Work. , in Australian Social Trends, March 2009. *Canberra: Australian Bureau of Statistics. Cat. No. 4102.0.*
5. Arulampalam, W. (2001). Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages. *The Economic Journal, 111(475), F585-F606*.
6. Billard, L. (2001). Markov Models and Social Analysis. In N. J. Smelser, P.B. Baltes (eds.), International Encyclopedia of the Social & Behavioral Sciences, pp. 9242 9250. Pergamon, Oxford.
7. Billari, F. C. (2001). Sequence Analysis in Demographic Research. *Canadian Studies in Population, Special Issue on Longitudinal Methodology* 28(2), 439- 458.
8. Burbidge, A. (2005). Australian Labour Market Transitions: Empirical Evidence. *University of Melbourne - No. CEDA paper: 82*.
9. Fasang, A. (2010), Retirement: Institutional Pathways and Individual Trajectories in Britain and Germany, *Sociological Research Online*, 15(2)
10. Fry J., and Boulton C. (2013), Prevalence of Transition Pathways in Australia. *Productivity Commission Staff Working Paper*.
11. Gabadinho, A., G. Ritschard, N. S. Mller and M. Studer (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software* 40(4), 1-37.
12. Green, C.P. and Leeves, G.D. (2009), Scarring, Habituation and Social Exclusion: Working Histories in Secure and Insecure Employment. *School of Economics Discussion Paper No. 385*. The University of Queensland
13. Hakim, C. (2000). Work-Lifestyle Choices in the 21st Century: Preference Theory. Oxford: Oxford University Press.
14. Haynes, M. Western, M., Yu, L. and Spallek, M. (2008). Analysing Nominal Data from a Panel Survey: Employment Transitions of Australian Women. Paper prepared for the 103rd Annual Meeting of the American Sociological Association, 1-4 August, 2008, Boston, USA.
15. Martin, R. and Morrison., P.S. (Eds.) (2001). *Geographies of labour market inequality*. TSO Publishing London.
16. McVicar D., and Anyadike, M. (2002). Predicting successful and unsuccessful transitions from school to work by using sequence methods. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 165(2), 317-334.
17. Saunders, P., and Brown, J. (2004). Explaining welfare to work transitions among the unemployed. *Australian Journal of Labour Economics*, 7(3), 395-410.
18. Stevens, A. H. (1997). Persistent Effects of Job Displacement: The Importance of Multiple Job Losses, *Journal of Labor Economics*, 15(1), 165-188
19. Studer, M., Ritschard, G., Gabadinho, A. & Mller, N.S. (2011), Discrepancy analysis of state sequences. *Sociological Methods and Research*, 40(3), 471-510.
20. Vickery, J. (1999). Unemployment and Skills in Australia *Reserve Bank of Australia (No. 1999-12)*.
21. Watson, N., and Wooden, M., (2002) The Household, Income and Labour Dynamics In Australia (HILDA) Survey. *HILDA Project Technical Paper Series. No, 1/02, May 2002.*
22. Wooldridge, J.M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA; The MIT Press.