1352.0.55.142

Australian Bureau of Statistics

Research Paper

Modelling the Short-Term Dynamics of Unemployment Using the ABS Longitudinal Labour Force Survey File

New Issue

Research Paper

Modelling the Short-Term Dynamics of Unemployment Using the ABS Longitudinal Labour Force Survey File

Cristian Ionel Rotaru

Analytical Services Branch

Methodology Advisory Committee

5–6 June 2014, Canberra

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30 AM (CANBERRA TIME) FRI 05 SEP 2014

ABS Catalogue no. 1352.0.55.142

Produced by the Australian Bureau of Statistics

INQUIRIES

The ABS welcomes comments on the research presented in this paper.
For further information, please contact Mr Ruel Abello, Analytical Services Branch on Canberra (02) 6252 6307 or email <analytical.services@abs.gov.au>.

MODELLING THE SHORT-TERM DYNAMICS OF UNEMPLOYMENT USING THE ABS LONGITUDINAL LABOUR FORCE SURVEY FILE

Cristian Ionel Rotaru
Analytical Services Branch

QUESTIONS FOR THE COMMITTEE

1. What are the views of the committee members on the approach adopted in this paper, in particular, the construction of time intervals, the subsequent discrete duration analysis, and the incorporation of random effects in the modelling?
2. This paper made use of the ABS Longitudinal Labour Force Survey (LLFS) file to examine the transitions out of unemployment. Any views on some other important economic or policy-related topics that can be analysed using the survey file?
3. Any views on some other (longitudinal) modelling techniques that can be applied to this new data source?

CONTENTS

MODELLING THE SHORT-TERM DYNAMICS OF UNEMPLOYMENT USING THE ABS LONGITUDINAL LABOUR FORCE SURVEY FILE

Cristian Ionel Rotaru
Analytical Services Branch

ABSTRACT

What affects the probability that an individual who has just entered unemployment finds employment within a given timeframe? Does the probability of exiting unemployment depend on the length of the individual's unemployment spell?

This paper reflects on these questions and analyses the transitions from unemployment of those aged 20–65 years, over the 2008–2010 period. The analysis makes use of the ABS Longitudinal Labour Force Survey (LLFS) file – a dataset that includes households that were followed for eight consecutive months during the said period. This paper is the first longitudinal analysis conducted on the file.

Building on the job-search theoretical framework, the paper builds a model aimed at analysing the factors that influence transitions from unemployment. A range of methodological techniques are implemented, including the creation of time intervals and the subsequent discrete duration analysis; the adoption of the competing-risks framework, to account for the different forms of exits from unemployment; and the inclusion of random effects in modelling the unobserved heterogeneity.

1. INTRODUCTION

What affects the probability that an individual who has just entered unemployment finds (full- or part-time) employment within a given timeframe? Does the probability of exiting unemployment depend on the length of the individual's unemployment spell?

This paper reflects on these important questions and analyses the transitions from unemployment in Australia for those aged 20–65 years over a three-year timeframe, from the beginning of 2008 to the end of 2010. One main contribution of this paper is its use of a new and important longitudinal data source, namely, the ABS Longitudinal Labour Force Survey (LLFS) file. By having more than 1.8 million records and around 150,000 households observed on a monthly basis, for a period of up to eight months, the dataset is well-suited for short-term dynamics of unemployment analyses. The sample also covers a period of considerable interest, the Global Financial Crisis (GFC).

In order to account for the discrete nature of the duration data – the data being collected on a monthly basis – discrete duration models are adopted in a competing-risks framework. In particular, three types of exits from unemployment are considered. The first is when the unemployed individual gets employed on a full-time basis (denoted by "FT"). The second is when the individual gets employed on a part-time basis (denoted by "PT"). The third one is when the individual leaves the labour force (denoted by "OLF"). As indicated in Flinn and Heckman (1983) the alternatives could be behaviourally different market states and as such it is important to treat them separately.

The analysis is divided into two parts. The first is focused on non-parametric techniques, and includes raw hazard and survival functions. The second incorporates observed as well as unobserved heterogeneity in modelling the hazard function. This is done using discrete duration models, where both the ordinary logit as well as the random effects logit models are examined.

The plan of the paper is as follows. Section 2 provides a conceptual background. Section 3 describes the data. The methodology used in the analysis is described in Section 4. Section 5 presents the results. Section 6 concludes.

2. CONCEPTUAL BACKGROUND

This paper makes use of the job-search theoretical framework (see Mortensen, 1970 and Lippman and McCall, 1976) to analyse the factors that affect the duration of unemployment. To see how the model works, consider an individual who has just become unemployed. This could happen either because the individual has moved from employment to unemployment or because he has transitioned from being out of the labour force to an active state of searching for work. Assuming that the individual continues to search for work until he gets employed, the aim is to explain the factors that determine the expected duration of remaining in the current state of unemployment. This expected duration is assumed to be inversely proportional to the probability of moving from unemployment to employment, which in turn is assumed to depend on two essential aspects:

1. the probability of receiving a job offer, and
2. the probability of accepting the job offer conditional on having received the job.

The probability of receiving a job offer is determined, amongst other things, by the demand for the individual's labour in the current market (Holzer, 1986). Amongst the things that employees look for are education, skills, and experience – factors that are aimed to make the individual more attractive to potential employers. Other factors that shift the demand include the local demand conditions, such as the business cycle and the phase of the local economy (e.g. the relative strength of the local economy and whether the economy is in recession) (Foley, 1997), as well as the search intensity of the individual (Holzer, 1986).

The probability of accepting the offer depends on the individual's reservation wage. This is the lowest wage at which the individual will accept a job offer. This wage depends on such factors as the expected wage in their particular occupation, family composition, other incomes in the household, unemployment benefits, as well as the probability of receiving future job offers and the expected work horizon (Long, 2009). Note that the reservation wage could also depend on some (or all) the determinants of the demand for the labour provided by the individual (Holzer, 1986).

From an econometrics perspective it is important to consider both aspects when analysing the determinants of unemployment duration. Failure to include one aspect might result in missing an essential component of the model which in turn could impact on the results. Note also that although the job-search model sets the theoretical framework for analysing the unemployment duration, empirical intervention is often needed to determine the effects (or the net effects) of the factors in the model.

As an example of the application of the model, consider the effects of the length of unemployment spell on the probability of exiting unemployment. On one hand, a longer unemployment spell could have negative consequences on the individual's prospects of finding work (Tansel and Tasci, 2010). One reason for this is the lack of investment in human capital due to the loss of valuable work experience during the unemployment spell. Another reason is the potential change in attitude, as the repeated failure to secure a job might discourage the individual from fully-exercising his skills in finding work (Foley, 1997). Finally, employers may be more reluctant to offer job offers to those with long unemployment spells. This is because they may perceive the long spell of unemployment as a signal of low productivity (Kroft *et al.*, 2013). These reasons are associated with a lower probability of receiving a job offer.

On the other hand, the individual might decrease his reservation wage as he gets closer to the end of his finite time horizon, so as to increase his prospects of securing a job (Lippman and McCall, 1976). This in turn will increase the conditional probability of accepting an offer. As the two effects move in opposite directions, it is not clear from theory which of them dominates. Empirical application would be useful to settle this uncertainty.

3. DATA AND DEFINITIONS

In this study, unemployment is defined as:

> Persons aged 15 years and over who were not employed during the reference week, and
>
> - had actively looked for full-time or part-time work at any time in the four weeks up to the end of the reference week and were available for work in the reference week; or
> - were waiting to start a new job within four weeks from the end of the reference week and could have started in the reference week if the job had been available then.
>
> (ABS, 2013)

The study uses data from the recently constructed ABS LLFS file that collects monthly information over a period of three years, from 2008 to 2010. The LLFS is compiled from 56 separate household surveys and records information about the labour market participation and employment transitions for all individuals in a household, over a period of up to eight consecutive months. Those in the scope of the survey are 15 years of age or older.

One main advantage of using the LLFS over other datasets is its wealth of information – the file includes more than 150,000 households and more than 1.8 million records. By recording monthly data for such a large number of households, over a period of up to eight consecutive months, the file is well-suited for in-depth analyses of short-term labour market dynamics.

For the purposes of this study, a number of restrictions were imposed. First, the sample was restricted to individuals who were between 20 to 65 years of age at the time of the first interview. Those older than 65 years or younger than 20 years were not included. Second, only private dwellings were included. Both these restrictions were imposed due to the potential different labour market behaviour of the individuals in these groups. Third, due to the aim of the study to analyse the duration of unemployment, the sample was further restricted to those who experienced unemployment at least once during the interview period. Also note that similar to Foley (1997), for the persons who experienced more than one spell of unemployment, only the first spell of unemployment was considered. This approach avoids the serial correlation that could result otherwise. As an extension, one could use multiple spells in the analysis by treating them as separate records. One would then need to deal with the dependence across spells.[1]

Note also that the analysis is restricted to the individuals who became unemployed during the interview period, i.e., restricted to the inflows in unemployment. This avoids the complexity of dealing with left truncation and potentially left-censoring.[2]

---

1 The methods included in Rotaru (2013) could be implemented to deal with this type of dependence.

2 Lancaster (1990) includes some stock sample techniques to deal with left-censoring and left-truncation.

Table A.5, in the Appendix, compares the sample included in the analysis to that excluded from the analysis by key covariates. Overall, the results look similar and there does not seem to be notable differences between the distributions of the two samples.

Table 3.1 below shows the distribution of the different types of exits from unemployment. Around 47% of the unemployed, in scope of this analysis, end in employment, of which around half end in full-time employment and half in part-time employment. Around 37% of the unemployed exit by leaving the labour force, a proportion which is similar to what other studies have found (see, for example, Morrison and Berezovsky, 2001). The balance of 16.5% remains in unemployment.

Compared to females, a higher proportion of males end in employment and a substantially higher percentage end in full-time employment. Females, on the other hand, are more likely to exit the labour force. In terms of marital status, the results are not too different across the two groups.

3.1 Percentage distribution of the exit states from unemployment spells

	Sex		*Marital status*		
	Male	*Female*	*Married*	*Not married*	*All*
Exiting unemployment via:					
Full-time employment	31.8	15.1	14.2	16.2	23.3
Part-time employment	19.5	27.6	29.1	25.8	23.6
Full- or part-time employment	51.3	42.7	43.3	41.9	46.9
Leaving the Labour Force (OLF)	31.2	41.8	43.4	39.9	36.6
Remaining in unemployment	17.5	15.5	13.3	18.2	16.5

4. METHODOLOGY

This paper models the transitions from the first unemployment spell of individuals observed during the eight consecutive months period described in Section 3. The aims are first, to model the probability of exiting unemployment and second, to account for the potential time dependence in the modelling.

In order to meet these aims and to adequately deal with the particulars of the duration data – the data for each household in the survey being collected on a monthly basis, for a period of up to eight waves – a few challenges need to be addressed. The first challenge is dealing with left-censored/truncated duration data, as some individuals were already unemployed at the time of the first interview. Although there are ways of addressing left-censoring or left-truncation (see Lancaster, 1990) the methods are considerably more complex and for left-truncation they rely on retrospective data, which might suffer from recall bias. To avoid this problem the analysis instead focuses on those who entered unemployment during the interview period.

The second challenge is dealing with the spells of individuals who have not yet exited unemployment at the time of the last interview and with the discrete nature of the data, the data for each individual being collected on a monthly basis for a period of up to eight months. To address these aspects of the data, the paper constructs time intervals and implements discrete duration modelling techniques. As these techniques require a more thorough exposition, they are elaborated more fully below.

The third challenge is controlling for the effects of covariates that are not available in the dataset, such as motivation and ability. This is addressed by including random effects in the modelling.

The fourth challenge is accounting for the different ways of exiting unemployment. To address this, the paper adopts the competing-risks framework.

Finally, the fifth challenge is dealing with the longitudinal aspect of the constructed person-period dataset, in which each individual has multiple records, one for each period. For this challenge, it is important to note that as explained later on in the section, by using the maximum likelihood estimation, the likelihood becomes a product of Bernoulli functions, which leads to simple estimation techniques. This means that relatively simple techniques are needed to estimate the parameters and that the well-known inferential statistics can be applied in this case. (See Muthén and Masyn, 2005; Singer and Willett, 1993 and Singer and Willett, 2003.)

4.1 Setting the framework

In a general setting consider a random sample of N unemployed individuals that are observed over a period of time, which in the context of this study is up to eight months long. This period is allowed to vary across individuals and is recorded on a discrete scale. The aim is to track the individuals from the time they entered into unemployment until they first exit that state (i.e., the focus is on single spells) and to analyse the characteristics that contribute to the differences in the duration experienced by the units in the sample.

The observed characteristics are captured in the vector $(X_1', X_2')'$,where in the context of the job-search model presented in Section 2, the first set of characteristics, assembled in vector X_1, determine the probability of being offered a job, whereas the second set, assembled in vector X_2, influence the probability of accepting the job offer. For example, X_1 could include sex, marital status, and education, whereas, X_2 could include family composition and marital status. Note that the two sets need not be mutually exclusive. As emphasised in Section 2, it is important to control for both sets of characteristics in the model. To further simplify the notation, all factors are collapsed into vector $X = (X_1', X_2')'$.

At the end of the spell, each unit i in the sample is assumed to end up in one of four states: exits unemployment and becomes employed full-time $(z_i = 1)$; exits unemployment and becomes employed part-time $(z_i = 2)$; remains unemployed $(z_i = 3)$ – case when the observation is censored; or exits the labour force altogether $(z_i = 4)$. Where here, as well as in the rest of the section, subscript i denotes the values for individual i.

To simplify the exposition consider the case where there is only one exit state, i.e., only one destination, case when the values of z are collapsed into a binary variable y, where $y_i = 1$ if individual i exits unemployment, i.e., when $z_i \in \{1,2,4\}$, and $y_i = 0$ if the individual remains in the initial state of unemployment, i.e., when $z_i = 3$. Note that with more than one exit state one can use the competing-risks framework presented in, for example, Singer and Willett (2003) and Allison (2010).

Let T_i^* be a random variable capturing the duration of unemployment for individual i, i.e., the duration until $y_i = 1$, and let $(0, t_{n_i}] = (0, t_0] \cup (t_0, t_{n_i}]$ be the interval over which the individual is observed. As the analysis is restricted to the inflows into unemployment, $s_0 := (0, t_0]$ is the interval in which the individual becomes unemployed. Further, t_{n_i} captures the last time the individual's responses are recorded (for the censored cases) or the first time the individual is known to have exited unemployment. n_i is the number of waves until individual i exits unemployment or until censoring after becoming unemployed. It follows that for censored cases T_i^* is not observed and all that is known is that $T_i^* > t_n$. Further, with the current dataset, even when $y_i = 1$, one does not necessarily know the exact T_i^*, and rather only knows that $T_i^* \in (t_{n_i - 1}, t_{n_i}]$. This is because the duration of unemployment is recorded on a monthly basis.

Although the approach taken by most empirical studies is to model the exact timing of event occurrence and treat duration as a continuous random variable, this paper takes a different path and instead models the probability that T^* falls into discrete time-intervals.

To see how this works, consider again the interval $(0, t_{n_i}] = (0, t_0] \cup (t_0, t_{n_i}]$ over which individual i is observed. The main idea is to transform the continuous time horizon into a sequence of discrete intervals. Now, partition the interval where the individual is "at risk" of leaving unemployment into n_i adjacent and mutually exclusive intervals, called periods, and which in the context of this study correspond to the time period between two consecutive waves of the survey, such that

$$(t_0, t_{n_i}] = (t_0, t_1] \cup (t_1, t_2] \cup \ldots \cup (t_{n_i - 1}, t_{n_i}]$$

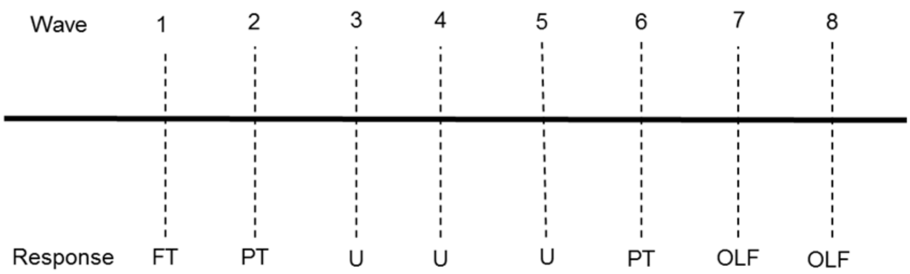
As a hypothetical example, consider an individual who is interviewed for all eight waves of the survey and whose responses are given in figure 4.1. In particular, at the time of the first interview he is employed full-time, in the second he is employed on a part-time basis, then he becomes unemployed and he remains so until before his sixth interview, at which time he indicates that he is employed part-time. During the time of the final two interviews he is out of the labour force.

Figure 4.2 shows how the intervals/periods were constructed. Note first that the analysis is focused on the period that starts with the entrance into unemployment (i.e., the period between 0 and t_0) and ends with the time at which the individual exits unemployment (i.e., the period between t_2 and t_3). Note also that as the information about his labour force status is collected at the time of the interview, it is unclear where exactly the transition between the different states of labour force occurred. As an example, consider the first interval s_0, where the individual enters unemployment. As it is only known that he was employed part-time at time 0 and that by the time of the next interview he has entered unemployment, it is unclear where exactly in the interval s_0 he entered unemployment. Rather than pinpointing to the exact time, the analysis instead focuses on intervals or periods. Using the same approach, the next two periods (s_1 and s_2) are constructed during which he is still unemployed. Finally, the final interval s_3 indicates the period when the individual exits unemployment into part-time employment.

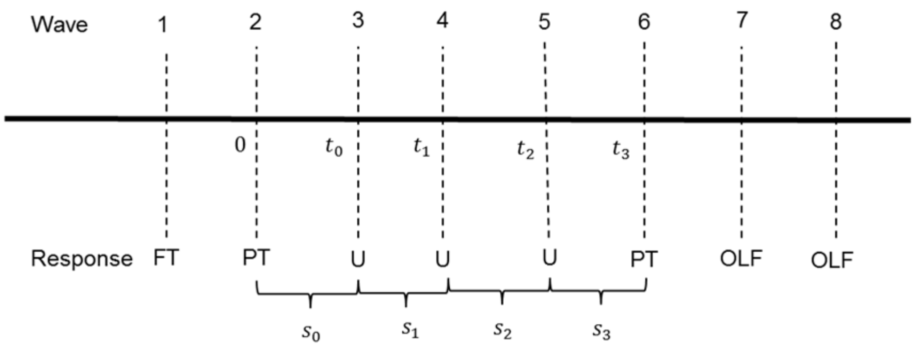
By disaggregating the duration into discrete time periods, one can proceed with the analysis by considering the discrete random variable T_i taking values from $\{1, 2, \ldots, n_i\}$, values which correspond to the n_i intervals, such that $T_i = j(i)$ whenever $T_i^* \in (\mathrm{t}_{j(i)-1}, t_{j(i)}] := s_{j(i)}$ for $j(i) = 1, 2, \ldots, n_i$. Note that hereafter, in order to simplify the notation, $j(i)$ will be replaced by j.

This strategy has therefore shifted the focus of the analysis from the continuous random variable T^* to the discrete-random variable T. This is appealing in this study because (1) there is a limited number of observations for each individual and (2) the exact timing of events is unknown. Although the timing of unemployment and the transition from this state might be captured by intervals, their exact timing is not covered in the data. By treating the duration using finite intervals, discrete survival models make adjustments for this limitation.

4.1 Example of a response



4.2 Example of a response – constructing the intervals



4.2 Modelling

The research question then analyses the duration T, given the observed characteristics X, the unobserved characteristics W, and the dependence over time, after controlling for the effect of censoring, which is captured by the indicator c. Mathematically, omitting the subscripts, the interest lies in mapping T to X, accounting for the time dependence (captured by t), potentially accounting for the unobserved characteristics W, and controlling for the effects of the censoring indicator c. (The variable W is dropped if the unobserved effects are not included in the analysis.) This is written as:

$$T = f_c(X, W, t) \tag{4.1}$$

where the subscript c in $f_c(.)$ indicates that the analysis controls for the effect of censoring.

Since T is by assumption intrinsically conditional (as it is assumed that individuals experiencing the target event have not experienced it before), interest lies in deriving its conditional probability function. The hazard function, which is well-suited and is central at analysing duration data, can be used for this scope.

In a discrete-time context, the hazard, h_{ij}, is defined as the conditional probability that individual i exits the state of unemployment in period j, which corresponds to interval $(t_{j-1}, t_j]$, given that the event has not occurred prior to period j.

Mathematically this is given by

$$h_{ij} := P\left(T_i = j \mid T_i \geq j\right) = P\left(t_{j-1} < T_i^* \leq t_j \mid T_i^* > t_{j-1}\right) \tag{4.2}$$

When covariates are included, one can extend (4.2) to the more informative hazard given by

$$h_{ij} := h_{ij}(x_{ij}, w_{ij}) := P\left(T_i = j \mid T_i \geq j, X_{ij} = x_{ij}, W_{ij} = w_{ij}\right) \tag{4.3}$$

where X_{ij} and W_{ij} are the vector of observed and respectively unobserved covariates for individual i and where x_{ij} and w_{ij} denote some particular values of X_{ij} and W_{ij}, respectively. Note that the hazard given in (4.3) is very flexible in that it includes covariates that are allowed to vary over time, as indicated by the subscript j. Note also that (4.3) conditions on a realised set of values of W_{ij}, although the covariates in W_{ij} are unobserved.

An important attractive feature of the hazard described in (4.3) is that since T_i is discrete, the hazard is simply a propensity and thus, one can use discrete choice models to model the duration of unemployment. In particular the common logit (used in this paper), the probit, and the complementary log-log models can be used. Specifically, the conditional probability of exit can be modelled as

$$P\left(T_i = j \mid T_i \geq j, X_{ij} = x_{ij}, \phi_i\right) = F\left(\gamma_m + x_{ij}'\beta + \phi_i\right) \tag{4.4}$$

where γ_m is a polynomial that needs to be specified by the researcher and which captures the duration dependence across the periods in the dataset, $F(.)$ is the *cdf*, and ϕ_i is a random component with known distribution which controls for the effects of the unobserved covariates. For example, if the logistic distribution is imposed, after some simple mathematics, (4.3) and (4.4) lead to the conditional log-odds:

$$\log\left[\frac{h_{ij}}{1 - h_{ij}}\right] = \gamma_m + x_i'\beta + \phi_i \tag{4.5}$$

Note first that in this paper $\gamma_m = \alpha_1 D_1 + \ldots + \alpha_m D_m$, which is a complete general specification for time. Here $\alpha = (\alpha_1, \ldots, \alpha_m)'$ is a vector of coefficients to be estimated, D_i are period indicators with a value of 1 for period i and 0 otherwise $(i = 1, \ldots, m)$, and m is the number of risk periods in the dataset. Note also that by dropping ϕ_i from equation (4.5) one gets back to the standard logit model applied to discrete-duration data.

Another advantage of the discrete duration model is that when the maximum likelihood approach is used to estimate the parameters, the resulting likelihood function is equivalent to a product of independent Bernoulli likelihood functions. This equivalence follows from the fact that the likelihood is given by

$$L(\beta, \alpha) = \prod_{i=1}^{N} \prod_{j=1}^{k_i} (1 - h_{ij})^{y_{ij}} h_{ij}^{1-y_{ij}} \tag{4.6}$$

where k_i stands for the number of observations where individual i is at risk in the person-period dataset, $y_{ij} = 1$ if the target event for individual i occurs at time j and $y_{ij} = 0$ otherwise, and where, for simplicity, the interest lies in estimating the values of α and β that maximise the likelihood function, ignoring the distribution parameters of the unobserved covariates.

It immediately follows that (4.6) is the product of a sequence of $k = k_1 + \ldots + k_N$ independent Bernoulli likelihood functions, which means that when conducting maximum likelihood estimations on the person-period dataset, the dichotomous y_{ij} values can be treated as if they were independent.

5. MODEL APPLICATION

One of the main aims of this analysis is to investigate the factors that affect the probability of exiting unemployment. As there are different exit states, the analysis uses the competing-risks framework to examine the duration of unemployment. Under this framework, one important assumption is that after conditioning on the regressors included in the model, the occurrence of each of the three events is non-informative for all the other states. This assumption allows for relatively simple parallel analyses where the analysis for each exit state is conducted on the same person-period dataset and where adjustments are only made to the censoring variable – i.e., treating the competing events as censored.

This section has two parts. The first presents the non-parametric results of unemployment duration analysis. Included are raw hazard rates for the whole sample as well as for some key covariates and a life table. The second part focuses on the modelling results of the discrete hazard function presented in Section 4. Included are the ordinary and the random effects logit model results.

5.1 Non-parametric results

Before proceeding with the regression analysis, this section begins with some simple nonparametric plots – raw hazard and survival function plots – and a life table. The results are presented in table 5.1, figure 5.2, and figure 5.3.

The results in table 5.1 indicate that the proportion of individuals moving out of unemployment decreases the longer they are unemployed. In particular, the largest proportion of exits occurs during the first period after they have become unemployed (around 60%) and decreases thereafter for each of the consecutive periods.

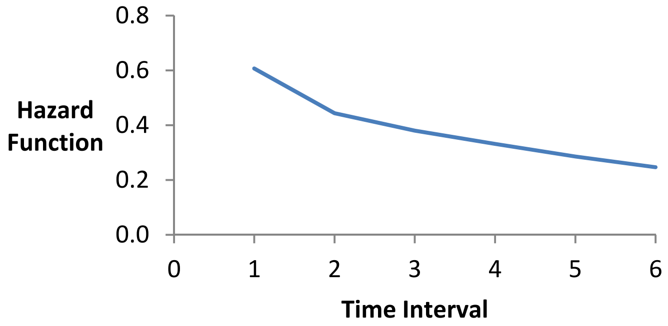
Figures 5.2 and 5.3 provide two graphical displays of this phenomenon. Figure 5.2, which displays the raw hazard rate function, shows that if the heterogeneity across individuals is ignored, the risk of leaving unemployment peaks during the first time interval and decreases thereafter, pretty sharply at first but quite smoothly after that. Figure 5.3, which shows the survival function, describes the same phenomenon – the survival function declining most sharply during the first time interval and decreasing at a decreasing rate thereafter. The results are similar to what Foley (1997) and Tansel and Tasci (2010) found.

To complement these results, Appendix A includes other exploratory results in the form of a life table for the different types of exits and hazard functions for sex and marital status. The results indicate that after ignoring the heterogeneity across individuals, the hazard rates tend to decrease with time for all exit types and that there seem to be differences in the hazard functions across both sex and marital status.

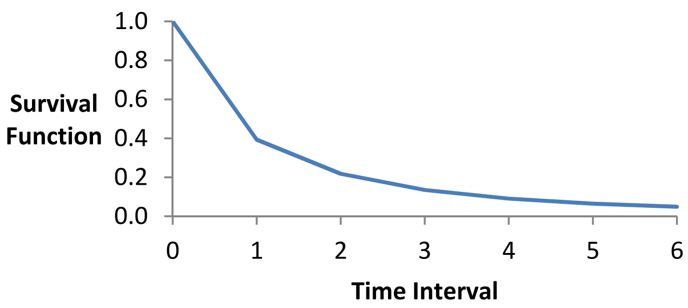
5.1 Life table describing the number of periods spent in unemployment

		Number of those:			*Proportion of those:*	
Period	*Time interval*	*Unemployed at the beginning of the period (Risk set)*	*Who left unemployment*	*Censored at the end of the period*	*Unemployed at the beginning of the period who left at the end of the period (Hazard function)*	*Still unemployed at the end of the period (Survival function)*
0	[0,1)	11,073	–	–	–	1.000
1	[1,2)	11,073	6,719	854	0.607	0.393
2	[2,3)	3,500	1,553	401	0.444	0.219
3	[3,4)	1,546	588	216	0.380	0.136
4	[4,5)	742	246	150	0.332	0.091
5	[5,6)	346	99	85	0.286	0.065
6	[6,8)	162	40	122	0.247	0.049

5.2 Hazard function for the duration of unemployment



5.3 Survival function for the duration of unemployment



5.2 Modelling results

Table 5.4 reports the results of the Ordinary Logit model. After controlling for the effect of the other covariates, the individuals aged 25–34 years have significantly higher odds of exiting unemployment via full-time employment than all the other age groups. They are then followed by those aged 35–44 years, the youngest group (those aged 20–24 years), and by those aged 45–54 years – the difference between the odds of the last three groups just mentioned being relatively small. With the exception of the oldest group, those aged 20–24 have higher odds of exiting into part-time employment. The oldest group, on the other hand, have much lower odds of exiting into full-time employment and significantly higher odds of leaving the labour force. To put it in perspective, when compared to the odds of the youngest group, the oldest group have around 52% lower odds of exiting into full-time employment and 73% higher odds of exiting the labour force.

In terms of gender, being male increases the odds of exiting into full-time employment by 27%, whereas being female increases the odds of exiting into part-time employment by 40%. For males, being married increases the odds of exiting into full-time employment, but decreases the odds of leaving the labour force and moving into part-time employment.

Compared to those with only secondary school completed, having higher education (Bachelor or TAFE) increases the odds of exiting into full-time employment (by at least 23%) and decreases the odds of exiting the labour force (by at least 28%). These results support the findings of Carroll (2006).

When compared to couples with no children and no other dependents, couples with dependents (children or other dependents) have lower odds of exiting unemployment via full-time work and higher odds of exiting unemployment via part-time work or into the OLF status. Based on the magnitude of the estimated coefficients, one-parent families with children are associated with the lowest odds of exiting unemployment via full-time employment, followed by couples with children. One-parent families with children have also the second highest odds of exiting the labour force, surpassed only by one-parent families with no children under 15 years, but with other dependents.

Overall, with the exception of technicians and trade workers, professionals are associated with higher or at least similar odds of exiting unemployment into full-time employment. Those who have last worked more than two years ago or are looking for work for the first time have much lower odds of exiting into any type of employment and much higher odds of exiting the labour force.

The results differ by location, with the Northern Territory, the Australian Capital Territory, Western Australia and Queensland (in that order) being associated with the highest odds of exiting unemployment via full-time employment. For the other exit types however the differences in the odds across states are not too different. The results also indicate that the individuals who reside in capital cities have a higher probability of exiting via full-time employment than those from the balance of state/territory.

Those from a non-English speaking background are associated with lower odds of exiting unemployment via full-time employment and the odds are lower if the individual arrived recently (i.e., after 2001). These results support those found in Carroll (2006). Non-English speakers who arrived after 2001 are associated with higher odds of exiting unemployment via part-time employment or into the OLF status.

The results for the initial unemployment quarter variable suggest a potential Global Financial Crisis effect. This is indicated by the big change in the magnitude of the odds around the first quarter of 2009. In particular, the individuals who have entered unemployment during this quarter have much lower odds of exiting unemployment via full-time or part-time employment. Those individuals who have entered unemployment during the first two quarters of 2009 are also associated with the largest odds of exiting into the OLF status.

For the time interval, the baseline logit hazard function confirms the previous results, as it is generally decreasing over time for all exit types. This implies that the conditional probabilities of exiting unemployment as well as that of exiting into the OLF status are lower with a longer unemployment spell.

Table 5.5 shows the estimation results of the random effects (RE) logit model. The RE logit model was applied to account for unobserved covariates, like ability or the intensity of job search. Overall, the RE logit model results are similar to those of the ordinary probit model, with the likelihood ratio test rejecting the null hypothesis in favour of the random effects model only in the case of the exit into full-time employment. Note that for all exit types, the AIC and BIC results are very similar across the two (i.e., ordinary and random effects) models.

5.4 Results for the Ordinary Logit Model – Log of Hazard Ratios

Variables	*Full-time*	*Part-time*	*Out of LF*
Age group *(20–24 years)*			
25–34 years	0.225 ***	–0.307 ***	–0.041
35–44 years	0.042	–0.263 ***	–0.003
45–54 years	–0.078	–0.215 ***	0.042
55–65 years	–0.730 ***	0.037	0.548 ***
Sex *(Female)*			
Male	0.241 ***	–0.339 ***	–0.156 **
Marital status *(Not married)*			
Married	–0.208 **	0.223 **	0.155 *
Male × married	0.837 ***	–0.404 ***	–0.352 ***
Education *(Secondary completed)*			
Bachelor	0.217 **	0.077	–0.331 ***
TAFE	0.206 ***	0.024	–0.345 ***
Secondary not completed	0.137	–0.247 ***	–0.153 **
Missing	0.279 ***	–0.399 ***	–0.253 ***
Family composition *(Couple, no children, no dependents)*			
Couple, no children, other dependents	–0.370 ***	0.235 **	0.217 **
Couple, children ,other dependents	–0.446 ***	0.244 ***	0.371 ***
One parent, children, other dependents	–0.722 ***	0.024	0.404 ***
One parent, no children, other dependents	–0.284	0.059	0.483 ***
One parent, no children, no other dependents	–0.189	–0.157	0.163
Lone person	–0.206 *	–0.029	0.023
Others	–0.033	0.306 ***	0.005
Last occupation *(Professional)*			
Manager	–0.068	–0.530 ***	0.144
Technician	0.169 *	–0.391 ***	0.188 *
Community	–0.633 ***	0.242 ***	0.23 **
Clerical	0.017	–0.307 ***	0.100
Sales	–0.317 ***	–0.135	0.131
Operator	–0.046	–0.477 ***	0.067
Labourer	–0.527 ***	0.004	0.254 ***
Last worked more than two years ago	–3.086 ***	–1.871 ***	1.152 ***
First time looking for work	–2.322 ***	–1.806 ***	1.012 ***
Missing	–1.949 ***	–2.062 ***	2.939 ***
State *(New South Wales)*			
Victoria	–0.109	0.129 **	–0.011
Queensland	0.187 ***	0.015	–0.163 **
South Australia	–0.224 **	0.073	–0.044
Western Australia	0.296 ***	0.079	0.092
Tasmania	–0.268 **	0.156	0.018
Australian Capital Territory	0.375 ***	0.038	0.020
Northern Territory	0.711 ***	–0.056	0.042
Language spoken *(English)*			
Non–English	–0.261 ***	–0.098	0.060
Year of arrival *(Arrived before 2001)*			
After 2001	–0.222 **	0.201 **	0.186 **

5.4 Results for the Ordinary Logit Model – Log of Hazard Ratios (continued)

Variables	*Full-time*	*Part-time*	*Out of LF*
Initial unemployment quarter *(Quarter 1, 2008)*			
Quarter 2, 2008	–0.158	–0.122	0.137
Quarter 3, 2008	–0.159	–0.080	–0.033
Quarter 4, 2008	–0.178 **	–0.135	–0.070
Quarter 1, 2009	–0.490 ***	–0.421 ***	0.173 *
Quarter 2, 2009	–0.270 **	–0.199 *	0.197 *
Quarter 3, 2009	–0.207 *	–0.188	0.063
Quarter 4, 2009	–0.294 ***	–0.068	–0.001
Quarter 1, 2010	–0.429 ***	–0.177 *	–0.057
Quarter 2, 2010	–0.037	–0.260 **	0.055
Quarter 3, 2010	–0.329 **	–0.040	–0.030
Quarter 4, 2010	–0.114	0.046	–0.127
Time interval			
1	–1.100 ***	–0.893 ***	–1.933 ***
2	–1.552 ***	–1.318 ***	–2.054 ***
3	–1.718 ***	–1.542 ***	–2.083 ***
4	–1.919 ***	–1.753 ***	–2.163 ***
5	–1.907 ***	–1.993 ***	–2.363 ***
6	–3.918 ***	–1.296 ***	–2.503 ***
Area of usual residence (Balance of state/territory)			
Capital city	0.185 ***	–0.211 ***	–0.006
Log likelihood	–6,307.3	–6,676.4	–7,895.2
AIC	12,724.6	13,462.7	15,900.4
BIC	13,151.5	13,889.7	16,327.4
Observations (n)	17,369	17,369	17,369

Notes:

*** $p< 0.01$; ** $p< 0.05$; * $p<0.10$.

Reference category is in brackets. Robust standard errors were computed.

5.5 Results for the Random Effects Logit Model – Log of Hazard Ratios

Variables	*Full-time*	*Part-time*	*Out of LF*
Age group *(20–24 years)*			
25–34 years	0.260 ***	–0.349 ***	–0.043
35–44 years	0.052	–0.296 ***	–0.003
45–54 years	–0.086	–0.243 **	0.044
55–65 years	–0.840 ***	0.043	0.574 ***
Sex *(Female)*			
Male	0.282 ***	–0.386 ***	–0.164 **
Marital status *(Not married)*			
Married	–0.228 *	0.246 **	0.159 *
Male × married	0.964 ***	–0.448 ***	–0.362 ***
Education *(Secondary completed)*			
Bachelor	0.231 **	0.091	–0.346 ***
TAFE	0.225 **	0.035	–0.361 ***
Secondary not completed	0.151	–0.277 ***	–0.157 **
Missing	0.324 ***	–0.438 ***	–0.263 ***
Family composition *(Couple, no children, no dependents)*			
Couple, no children, other dependents	–0.440 ***	0.268 **	0.225 **
Couple, children ,other dependents	–0.529 ***	0.272 ***	0.387 ***
One parent, children, other dependents	–0.822 ***	0.015	0.422 ***
One parent, no children, other dependents	–0.320	0.069	0.501 ***
One parent, no children, no other dependents	–0.231 *	–0.182	0.169
Lone person	–0.222 *	–0.038	0.022
Others	–0.039	0.345 ***	0.003
Last occupation *(Professional)*			
Manager	–0.094	–0.608 ***	0.147
Technician	0.199 *	–0.447 ***	0.196 *
Community	–0.735 ***	0.305 **	0.239 **
Clerical	0.004	–0.354 ***	0.106
Sales	–0.377 ***	–0.153	0.132
Operator	–0.042	–0.551 ***	0.069
Labourer	–0.624 ***	0.012	0.262 ***
Last worked more than two years ago	–3.356 ***	–2.077 ***	1.215 ***
First time looking for work	–2.564 ***	–2.011 ***	1.058 ***
Missing	–2.151 ***	–2.233 ***	3.054 ***
State *(New South Wales)*			
Victoria	–0.123	0.151 *	–0.012
Queensland	0.227 ***	0.025	–0.173 **
South Australia	–0.260 **	0.077	–0.048
Western Australia	0.346 ***	0.092	0.092
Tasmania	–0.298 **	0.180	0.017
Australian Capital Territory	0.438 ***	0.051	0.022
Northern Territory	0.823 ***	–0.066	0.038
Language spoken *(English)*			
Non–English	–0.303 ***	–0.116	0.063
Year of arrival *(Arrived before 2001)*			
After 2001	–0.246 *	0.246 *	0.193 **

5.5 Results for the Random Effects Logit Model – Log of Hazard Ratios (continued)

Variables	*Full-time*	*Part-time*	*Out of LF*
Initial unemployment quarter *(Quarter 1, 2008)*			
Quarter 2, 2008	–0.194	–0.148	0.144
Quarter 3, 2008	–0.202	–0.096	–0.035
Quarter 4, 2008	–0.213 **	–0.148	–0.071
Quarter 1, 2009	–0.567 ***	–0.485 ***	0.182 *
Quarter 2, 2009	–0.315 **	–0.236 *	0.210 *
Quarter 3, 2009	–0.247 *	–0.234 *	0.069
Quarter 4, 2009	–0.351 ***	–0.084	–0.001
Quarter 1, 2010	–0.510 ***	–0.206 *	–0.058
Quarter 2, 2010	–0.072	–0.301 **	0.058
Quarter 3, 2010	–0.385 **	–0.058	–0.030
Quarter 4, 2010	–0.139	0.048	–0.133
Time interval			
1	–1.259 ***	–1.017 ***	–2.001 ***
2	–1.573 ***	–1.322 ***	–2.083 ***
3	–1.632 ***	–1.459 ***	–2.082 ***
4	–1.755 ***	–1.603 ***	–2.128 ***
5	–1.679 ***	–1.798 ***	–2.303 ***
6	–3.655 ***	–1.049 **	–2.411 ***
Area of usual residence *(Balance of state/territory)*			
Capital city	0.213 ***	–0.244 ***	–0.007
Log likelihood	–6,305.7	–6,675.7	–7,894.8
Sigma	0.887	0.851	0.413
Rho+	0.193 **	0.180	0.049
AIC	12,723.4	13,463.5	15,901.6
BIC	13,158.1	13,898.2	16,336.3
Observations (n)	17,369	17,369	17,369

Note:
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.
+ = likelihood ratio test for rho = 0.
Reference category is in brackets. Robust standard errors were computed.

6. CONCLUDING REMARKS

Building on the job-search theoretical framework, this paper examined the transitions from unemployment using the ABS Longitudinal Labour Force Survey (LLFS) file. The file covers a three-year period, from the beginning of 2008 to the end of 2010. By including more than 1.8 million records from around 150,000 households observed over a period of up to eight consecutive months, the file is useful for analysing short-term labour market dynamics.

From a methodological perspective, the paper implemented the following techniques to deal with the specific features of the data and of the analysis. First, to simplify the modelling, the analysis was restricted to those who were observed to have become unemployed during the eight-month interview period. This approach avoided the reliance on retrospective information and the model complexities involved with dealing with left censoring/truncation. Second, to capture the discrete nature of the duration data and to deal with left censoring, discrete duration models were implemented. This strategy shifted the focus of the analysis from modelling a continuous random duration variable to that of an analysis conducted on time intervals. Third, to consider the different unemployment exits, the analysis adopted the competing-risks framework and separately examined the transition into three different exit states: full-time employment, part-time employment, or out of the labour force. Finally, to account for unobserved heterogeneity, random effect models were also considered.

From an empirical perspective – and also in response to the two questions posed in the introduction – the following can be noted. First, the results differ by age groups with the older workers (aged 55–65 years) having significantly lower odds of exiting unemployment into full-time employment and with much higher odds of exiting the labour force. Second, when compared to the other types of families and after controlling for the effect of the other covariates, lone parents with children under 15 years have the lowest odds of exiting into full-time employment and the second highest odds of exiting the labour force, surpassed only by lone parents with other dependents, but with no children under 15 years. Third, the results also differ significantly by state with the Australian Capital Territory and three main mining states of Australia (Western Australia, Northern Territory, and Queensland) being associated with the highest odds of exiting unemployment into employment. Fourth, similar to the results of Carroll (2006), those from a non-English background have lower odds of exiting into full-time employment. Apart from these, the hazard function is also influenced by the other covariates, which include sex, marital status, education, occupation, initial unemployment quarter, and area of usual residence. Finally, the results indicate that the probability of exiting unemployment depends on the length of unemployment spell with the baseline hazard function decreasing over time.

ACKNOWLEDGEMENTS

The author is particularly thankful to Yovina Joymungul Poorun, Franklin Soriano, Ruel Abello, Dr. Siu-Ming Tam, Professor James Brown, and the ABS Methodology Advisory Committee members, for their valuable suggestions and comments.
An earlier version of this paper was presented at the Australian Labour Market Research Workshop in February 2014 and the author would like to thank the discussant, Dr. Marcia Keegan, and the workshop participants for their helpful comments. Thanks also go to Peter Rossiter for his help in editing and formatting the paper. Remaining errors are the author's.

REFERENCES

Allison, P.D. (2010) *Survival Analysis using SAS®: A Practical Guide*, Second edition, SAS Institute Inc., Cary NC.

Australian Bureau of Statistics (ABS) (2013) *Labour Force, Australia*, cat. no. 6202.0, ABS, Canberra

Carroll, N. (2006) "Explaining Unemployment Duration in Australia", *Economic Record,* 82(258), pp. 298–314.

Cupples L.A.; D'Agostino R.B.; Anderson K. and Kannel W.B. (1988) "Comparison of Baseline and Repeated Measure Covariate Techniques in the Framingham Heart Study", *Statistics in Medicine*; 7(1–2), pp. 205–218.

Flinn, C.J. and Heckman, J.J. (1983) "Are Unemployment and Out of the Labour Force Behaviourally Distinct Labour Market States?", *Journal of Labour Economics,* 1(1), pp. 28–42.

Foley, C.M. (1997) "Determinants of Unemployment Duration in Russia", Discussion Paper No. 779, Economic Growth Center, Yale University.

Holzer, H.J. (1986) "Black Youth Nonemployment: Duration and Job Search", in Freeman, R.B. and Holzer, H.J. (eds), *The Black Youth Employment Crisis*, University of Chicago Press, Chicago.

Kroft, K.; Lange, F. and Notowidigdo, M.J. (2013) "Duration Dependence and Labour Market Conditions: Evidence from a Field Experiment", *Quarterly Journal of Economics*, 128(3), pp. 1123–1167.

Lancaster, T. (1990) *The Econometric Analysis of Transition Data,* Cambridge University Press.

Lippman, S. and McCall, J. (1976) "The Economics of Job Search: A Survey (Parts I and II)", *Economic Inquiry*, 14(2), pp. 155–189 and 14(3), pp. 347–368.

Long, K. (2009) "Unemployment Durations: Evidence from the British Household Panel Survey", *Economic & Labour Market Review*, 3(10), pp. 48–54.

Morrison, P. and Berezovsky, O. (2001) "Labour Market Risk and the Regions: Evidence from the Gross Labour Flows", in Martin, R. and Morrison, P. (eds), *Geographies of Labour Market Inequality*, TSO Publishing, London.

Mortensen, D.T. (1970) "Job Search, the Duration of Unemployment, and the Phillips Curve", *American Economic Review,* 60(5), pp. 847–862.

Muthén, B. and Masyn, K. (2005) "Discrete-Time Survival Mixture Analysis", *Journal of Educational and Behavioural Statistics*, 30(1), pp. 27–58.

Rotaru, C.I. (2014) "Discrete Choice Panel Data Modelling using the ABS Business Longitudinal Database", *Methodology Advisory Committee Papers,* cat. no. 1352.0.55.139, Australian Bureau of Statistics, Canberra.

Singer, J.D. and Willett, J.B. (1993) "It's about Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events", *Journal of Educational Statistics,* 18(2), pp. 155–195.

Singer, J.D. and Willett, J.B. (2003) *Applied Longitudinal Data Analysis: Modelling Change and Event Occurrence*, Oxford University Press, New York.

Tansel, A. and Tasci, H.M. (2010) "Hazard Analysis of Unemployment Duration by Gender in a Developing Country: The Case of Turkey", *Labour,* 24(4), pp. 501–530.

APPENDIXES

A. ANALYSIS RESULTS

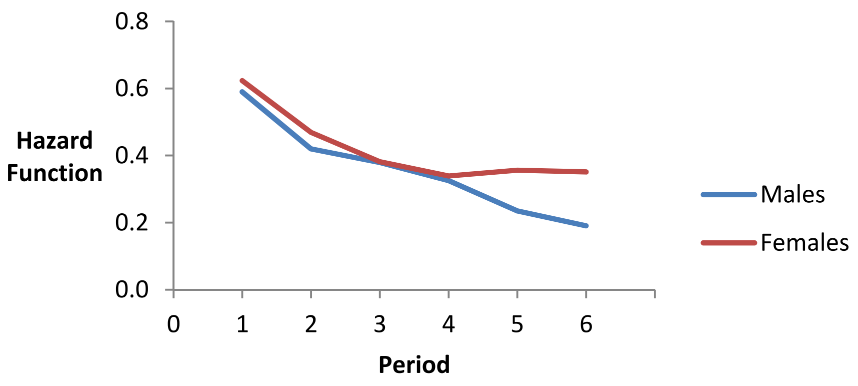
A.1 Hazard functions for exiting unemployment

Time period	*Any exit*	*Full-time*	*Part-time*	*Out of LF*
0	–	–	–	–
1	0.607	0.171	0.173	0.263
2	0.444	0.123	0.125	0.196
3	0.380	0.107	0.102	0.171
4	0.332	0.084	0.082	0.166
5	0.286	0.081	0.064	0.142
6	0.247	0.012	0.105	0.130

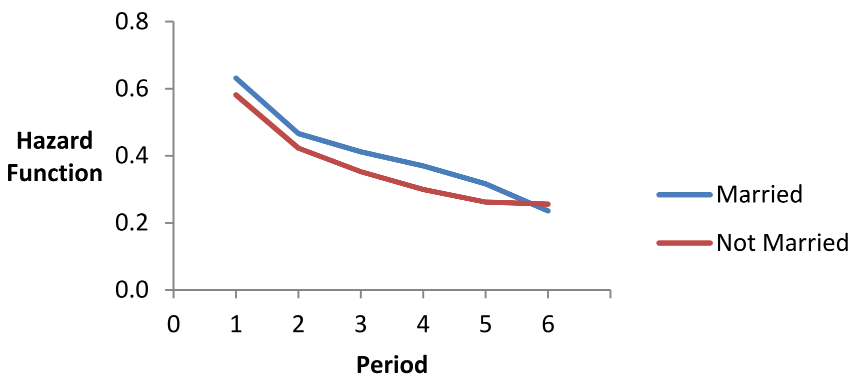
A.2 Survival functions for exiting unemployment

Time period	*Any exit*	*Full-time*	*Part-time*	*Out of LF*
0	1.000	1.000	1.000	1.000
1	0.393	0.829	0.827	0.737
2	0.219	0.727	0.723	0.593
3	0.136	0.650	0.649	0.491
4	0.091	0.595	0.596	0.410
5	0.065	0.547	0.558	0.352
6	0.049	0.540	0.500	0.306

A.3 Hazard function for the duration of unemployment, by Sex



A.4 Hazard function for the duration of unemployment, by Marital status



A.5 Comparison between the sample included in the analysis and that excluded from the analysis, by key covariates

Variable	*Included**	*Excluded***
	(%)	(%)
Sex		
Male	49.1	52.4
Female	50.9	47.6
Marital status		
Married	50.9	43.4
Not married	49.1	56.6
State		
New South Wales	23.7	27.5
Victoria	22.2	21.8
Queensland	17.9	17.5
South Australia	10.9	11.4
Western Australia	12.7	10.4
Tasmania	5.8	6.8
Northern Territory	3.4	2.3
Australian Capital Territory	3.4	2.3
Occupation		
Manager	4.9	3.7
Professional	11.2	8.4
Technician	10.8	8.3
Community	8.5	7.1
Clerical	10.5	9.3
Sales	7.5	6.5
Operator	5.9	6.3
Labourer	14.7	16.6
Last worked more than two years ago	8.0	14.1
First time looking for work	4.0	5.9
Missing	14.0	13.8
Education		
Bachelor	14.7	11.8
TAFE	25.2	20.3
Secondary completed	13.7	10.6
Secondary not completed	17.8	18.1
Missing	28.6	39.2
Family composition		
Couple, no children, no other dependents	29.4	26.9
Couple, no children, other dependents	7.0	5.9
Couple, children, other dependents	26.3	23.4
One parent, no children, no other dependents	5.2	7.8
One parent, no children, other dependents	2.4	2.5
One parent, children, other dependents	8.4	9.8
Lone person	9.9	12.0
Others	11.3	11.6
Age group		
20–24 years	20.9	20.6
25–34 years	26.6	26.6
35–44 years	23.0	22.5
45–54 years	17.4	19.5
55–65 years	12.1	10.8

Notes: * refers to the observations included in the analysis – those individuals that became unemployed during the interview period; ** refers to the observations excluded from the analysis – those individuals that were unemployed at the time of the first interview.

B. DATA COMPILATION

This section lists the variables used in the models.

State
- New South Wales
- Victoria
- Queensland
- South Australia
- Western Australia
- Tasmania
- Northern Territory
- Australian Capital Territory

Sex
- Male
- Female

Age group
- 20–24 years
- 25–34 years
- 35–44 years
- 45–54 years
- 55–65 years

Marital status
- Married
- Not married

Occupation
- Managers and administrators
- Professionals
- Technicians and trade workers
- Community and professional service workers
- Clerical and administrative workers
- Sales workers
- Machinery operators and drivers
- Labourers
- Last worked more than two years ago
- First time looking for work
- Missing

Education
- Degree – Bachelor or Postgraduate degree
- TAFE – Diploma or Certificate
- Secondary school completed
- Secondary school not completed

Language spoken
- English
- Non-English

Year of arrival in Australia (Non-English speakers)
- Arrived before 2001
- Arrived after 2001

Initial unemployment quarter
- Quarter 1, 2008
- Quarter 2, 2008
- Quarter 3, 2008
- Quarter 4, 2008
- Quarter 1, 2009
- Quarter 2, 2009
- Quarter 3, 2009
- Quarter 4, 2009
- Quarter 1, 2010
- Quarter 2, 2010
- Quarter 3, 2010
- Quarter 4, 2010

Family composition
- First digit: family
- Second digit: number of parents (1 – single and 2 – couple)
- Third digit: whether the family has children under 15
- Fourth digit: whether the family has other dependents
  - 0000 – Lone person
  - 1100 – One parent family with no children and no other dependents
  - 1101 – One parent family with no children under 15 and other dependents
  - 1111 – One parent family with children under 15 and other dependents
  - 1200 – Couple family with no children and no other dependents
  - 1201 – Couple family with no children under 15 and other dependents
  - 1211 – Couple family with children under 15 and other dependents
  - 9999 – Others

Time interval
- This splits the period of 8 waves into intervals
  - 1
  - 2
  - 3
  - 4
  - 5
  - 6 (i.e. the last two periods were combined because of the small sample sizes)

Area of usual residence
- Capital city
- Balance of state/territory

FOR MORE INFORMATION . . .

INTERNET	**www.abs.gov.au** The ABS website is the best place for data from our publications and information about the ABS.
LIBRARY	A range of ABS publications are available from public and tertiary libraries Australia wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our website for a list of libraries.

INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our website, or purchase a hard copy publication. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

PHONE	1300 135 070
EMAIL	client.services@abs.gov.au
FAX	1300 135 211
POST	Client Services, ABS, GPO Box 796, Sydney NSW 2001

FREE ACCESS TO STATISTICS

All statistics on the ABS website can be downloaded free of charge.

WEB ADDRESS	www.abs.gov.au