The Australian National University

Centre for Economic Policy Research

DISCUSSION PAPER

Explaining Unemployment Duration in Australia

Nick Carroll

DISCUSSION PAPER NO. 483

December 2004

ISSN: 1442-8636
ISBN: 0 7315 3553 7

PhD Scholar, Economics Program, Research School of Social Sciences, Australian National University, Canberra, ACT 0200 Australia. Email: Nick.Carroll@anu.edu.au, Tel: +61 2 6125 3854

Acknowledgements
I am grateful for assistance from Alison Booth, Gary Barrett and Tue Gorgens and for the helpful comments received from participants at a workshop at RSSS in July 2004 and at the ANU/WA PhD conference in November 2004. The usual disclaimers apply.

CONTENTS

ABSTRACT

What influences the probability that someone will leave unemployment? Informed by a search-theoretic framework and allowing for exits to not in the labour force and employment, I examine what influences the probability that somebody will leave unemployment. The unemployment data used are derived from the retrospective work history information from the first two waves of the Household, Income and Labour Dynamics in Australia Survey. I find that variables that increase wage offers and lower reservation wages are associated with shorter durations of unemployment and that exit rates from unemployment appear to remain steady initially with duration before declining relatively sharply.

Keywords: survival analysis unemployment durations

JEL Classification: J64

1. Introduction

What influences the probability that someone will leave unemployment? Does the probability of exit from unemployment fall the longer somebody has been in unemployment? Is the negative relationship commonly observed between length of spell and exit probability related to "low exiters" remaining in unemployment, or due to unemployment being harmful for your labour market health?

A number of authors have found that unemployment affects people's happiness and life-satisfaction (Winkleman and Winkleman (1998) and Clark and Oswald (1994)), wages (Ruhm (1991)[1] and Arulampalam (2001)), and probability of unemployment in future periods (Arulmpalam, Booth and Taylor (2000)). In addition, if unemployment is scarring we may potentially see exit rates fall as the current spell of unemployment lengthens (Machin and Manning (1999)). While in this literature there are issues about whether unemployment causes worse outcomes (state dependence), or whether people with certain characteristics become unemployed (unobserved heterogeneity), there appears to be a general finding that unemployment has negative impacts on the people that experience it.

While unemployment appears to be associated with worse outcomes, there is a wide variation in the unemployment experience. Some unemployment experiences are short and singular, while other groups of people have repeated and longer spells of unemployment. Owing to data limitations, research in Australia into unemployment spells has tended to be quite descriptive in nature (Dockery (2003)) or focus on demographic characteristics and limited labour market and personal characteristics (see Borland (2000), Chapman and Smith (1992), Brooks and Volker (1986)).

This paper investigates the variables associated with exits from unemployment using a duration modelling framework. The analysis enables us to see what characteristics are associated with early exit from unemployment and to investigate, holding certain observable characteristics constant, how exit rates vary with duration. This paper makes three contributions to the literature. Firstly, it uses an important new longitudinal data source (Household, Income and Labour Dynamics in Australia (or HILDA) Survey) from a recent period to investigate the factors associated with exit from unemployment. In so doing it uses more flexible baseline hazards than have typically been used in unemployment duration modelling in Australia.

[1] Ruhm (1991) specifically examined the impact of layoffs on future wages.

Secondly, the paper examines several factors associated with unemployment exit that have not previously been examined in Australian research. In particular, employment experience, unemployment experience and risk aversion are included in estimation. Risk aversion is potentially important, given the job-seekers' choice about whether to accept a certain wage offer in the current period (hence exiting unemployment) or wait for an inherently uncertain wage in a future period that may be higher (hence lengthening duration). The third key insight that this paper offers is a discussion of a flexible underlying baseline hazard. Australian research in this area has typically assumed that changes in the hazard rate are monotonic and specifications have not been flexible enough to allow for non-monotonicity.

The first result from this paper is that factors expected to raise a person's wage offers (such as employment experience) and lower their reservation wage (such as ineligibility for some benefits) are associated with higher exit rates from unemployment. The second result is that exit rates from unemployment do not fall with duration initially, but after a period of approximately four months they then appear to decline by a large amount. Because of this, modelling that assumes a monotonic relationship between spell length and exit rates will tend to under-estimate the decline in the hazard at longer durations[2] (because it will be a weighted average of the declining and non-declining components). The final key result is that, even with the number of variables used in the estimation presented here, they 'explain' less than one fifth of the decline in the baseline hazard. This suggests that either unobserved characteristics are potentially important or that some scarring is occurring.

The remainder of this paper is structured as follows. Section II outlines a general search framework used for analysis, section III briefly describes the econometric tools, section IV describes the data used, section V gives the results from the full model and section VI presents two extended models focusing on the role of risk and previous wages, section VII examines the baseline hazard and section VIII provides conclusions. In Appendix I descriptions of the variables are provided and a supporting paper (Carroll (2004)) provides more detailed explanation of the data and methods.

[2] It should also be noted that the coefficients will also be affected by mis-specifying the baseline hazard and it may be reflected in the violation of the proportional hazards assumption.

2. The Search Framework

Because my data set is primarily a person level data set I select and interpret my variables using the standard search theoretic framework (see Mortenson (1986)). With a simple search theoretic model, a person who is searching for a job receives one offer per period[3] of x from a wage distribution given by $F_n(x)$. There is a search cost, c, a value of not working (leisure, home production, or the unemployment benefit), b, and a discount rate B. If the searcher accepts the wage offer, x, they receive this for n periods, if they decline the wage offer then in the next period they draw another wage.

I denote $V_n(x)$ as the maximal B-discounted expected return attainable when n periods remain and the currently available job offer is x (with future wages appropriately discounted). The searcher will choose the state that gives the maximum return, i.e. the wage with the current job offer or continued search (which includes the expected returns from future search). Thus, $V_n(x)$ satisfies the recursive equation (i.e. V_n is the period prior to V_{n-1}):

$$V_n(x) = \max \{ x, b\text{-}c + B \int_0^\infty V_{n-1}(y) \, \partial F_{n-1}(y) \} \tag{1}$$

where x is the draw at time n and y is the draw at time $n\text{-}1$ and where the integral is over 0 to ∞ because we are summing up over all possible wage offers at time $n\text{-}1$.

As shown in the search literature,[4] the workers best strategy is to choose a reservation wage that maximises expected utility. The worker will then accept a job if the wage is higher than the reservation wage. That is the searcher will accept the wage if the returns from accepting the wage are greater than the expected returns from declining the wage offer. Thus, the reservation wage at time n (denoted by R_n) will be set equal to the expected net benefits of future search thus:

$$R_n = b\text{-}c + B \int_0^\infty V_{n-1}(y) \, \partial F_{n-1}(y) \tag{2}$$

The exit rate from unemployment in each period is given by $(1 - F_n(R))$, where $F_n(R)$ is the probability that a wage drawn from wage distribution $F_n()$ will be below a person's reservation wage, R. Where the arrival rate varies from 1 per period, the expected exit rate per period is then given by the probability of an offer arriving (denoted

[3] This assumption is made to simplify the discussion. Where arrival rates are allowed to vary (i.e. the number of offers can vary between 0 and infinity) the returns to future search is the expected returns from all possible arrival rates multiplied by their probability of occurring.
[4] See Mortensen (1986).

by τ) multiplied by the probability of the offer being accepted, thus the exit rate is $\tau (1-F_n(R))$. The probability that a person will exit unemployment is a function of the rate at which offers arrive, the wage distribution that they draw from, and their reservation wage. The expected duration of unemployment is then given by the inverse of the exit rate $(1 / \tau (1- F_n(R)))$.

Thus, the higher the wage offers relative to the reservation wage the higher the exit rate, and alternatively, the lower the reservation wage relative to the wage offers the higher the exit rate. Higher non-wage income will result in a higher reservation wage (because the benefits of future search are higher), resulting in a lower probability of a wage offer being above the reservation wage and hence a lower exit rate. Higher wage offers, all else constant,[5] will result in a higher probability of a wage offer being accepted and therefore a higher exit rate.

3. Estimation of Survival Time Models

I use survival data (where a spell is observed over time until an event occurs). In the case of unemployment we observe a person enter unemployment and we observe the spell until exit. Two key concepts are the hazard rate (the proportion of an 'at-risk' group that leaves a particular state over a reference period) and the survival rate (the proportion of the initial group that remains in the state at the reference period).

I use a proportional hazards model, whereby explanatory variables move a baseline hazard up and down by a fixed proportion. The proportional hazard model takes the following form:

$$h_i(t)=h_0 (t) \exp (z_i(t)' \beta) \qquad (3)$$

where $h_i(t)$ is the hazard for person i, $h_0(t)$ is the common baseline hazard, $z_i(t)$ are the observable characteristics, and β and h_0 are the parameters to be estimated.

I use two estimation methods with different assumptions about the baseline hazard to produce the results below. In the first estimation I assume that there is a common baseline hazard to all people, but I do not restrict the shape of the baseline hazard to be of a certain shape (semi-parametric estimation). In the second estimation I

[5] Higher expected future wages will result in increases in reservation wages, but not by the full amount of the increase in the expected future wage. This effect is taken into account with the statement 'all else constant'.

divide the time axis into a finite number of intervals and estimate a separate baseline hazard parameter for each interval.[6] The approaches that I use are very flexible methods for estimating the baseline hazard, which avoid the problems associated with imposing a parametric functional form.[7] I use the standard partial log likelihood function for my semi-parametric estimation (see Lancaster (1990) for more details about the log likelihood functions given in this section):

$$L(\beta) = \sum_{j=1}^{D} [\sum_{m \in D_j} z_m \beta - d_j \log \{\sum_{i \in R_j} \exp (z_i \beta)\}] \qquad (4)$$

where j indexes the ordered failure times t_j, D_j is the set of d_j observations that fail at t_j, d_j is the number of failures at t_j and R_j is the set of observations m that are at risk at time t_j. I use the standard likelihood function for the piecewise constant estimation with N unemployment spells:

$$\mathrm{L}(\beta,\gamma) = \sum_{\mathrm{i=1}}^{\mathrm{N}} \delta_i \log(1 - \exp(-\exp[\gamma(k_i) + z_i (k_i)'\beta])) - \sum_{\mathrm{t=0}}^{\mathrm{k_i-1}} \exp[\gamma(t) + z_i(t)'\beta] \qquad (5)$$

where k_i is the observed length of the ith spell, δ_i=1 if the spell is not right censored (0 otherwise). In maximising the log likelihood the $\gamma(t) = \ln[\int_t^{t+1} h_0(u).du]$ are treated as parameters to be estimated.

I use the independent competing risks framework[8] to examine exits to employment separately from exits to not in the labour force. We assume that there are two latent survival times, one for employment and one for not in the labour force, and the actual destination observed is the minimum of the latent survival times. With this assumption, exits from unemployment to not in the labour force are treated as censored spells for exits to employment in estimation.[9] We then use the likelihood function in equations (4) and (5) for estimation. In estimation below I also allow for unobserved

[6] The intervals (in thirds of a month) that I employ are: 1,2,3,4,5,6,7-9,10-12, 13-15,16-18, 19-24, 25-30, 31-36, 37+.

[7] Barrett (2000) (drawing from the literature) notes that misspecification of the baseline hazard is a major source of error in drawing inferences concerning both the presence of duration dependence and the impact of covariates.

[8] A key assumption of the independent competing risks framework is that the risk of exits to employment and not in the labour force are independent, *conditional on the covariates*.

[9] In the discrete setting it is possible to undertake estimation with a variety of assumptions about how risk of exit may vary within discrete periods (see Narendrenathan and Stewart (1993)). However, these assumptions are beyond the scope of this paper.

individual level heterogeneity. I use the method suggested by Heckman and Singer (1984), whereby the unobserved heterogeneity distribution is estimated non-parametrically by a discrete multinomial distribution.

4. Household, Income and Labour Dynamics in Australia (HILDA) Survey

The data used for this study are from the HILDA panel database. The survey is primarily administered in the second half of each year, with the first wave being collected in the second half of 2001. Currently two waves of data are available (2001 and 2002) and it is expected that the third wave will become available in early 2005. In wave 1, 7682 households were sampled comprising 13,969 members. The HILDA survey has a wide variety of variables on health, family background, work history, demographics and educational and training history, which allows us to control for some heterogeneity.

The data used for the analysis of spells in the HILDA database is drawn from a calendar (see Carroll (2004)). In this calendar the respondent is asked to provide their employment status in thirds of a month over the past 12-18 months (depending on when the interview takes place). The unit of analysis for the remainder of the paper is thirds of a month. From table 1 we can see that we have 2402 spells of unemployment, with 1542 exits to employment or not in the labour force (referred to as fails) and 860 spells where the end of the spell is not observed in the sample period (in other words the spell is right censored[10]). Out of the 2402 spells, 1757 of these spells begin during the sample period, 471 spells are left censored (for a description see sub-section (ii) of section IV) and 174 spells are left truncated (also see sub-section (ii) of section IV).

Table 1: what proportions of spells are affected by left censoring/ left truncation (unit: spells)

	No fail	Fail	Total
Left truncated	122	52	174
Left censored	92	379	471
Begin in scope	646	1,111	1,757
Total	860	1,542	2,402

Figure 1 presents the raw hazard rates. We see the general decline in the hazard rates over time, from 6-9% at the start of the spell to 2-5% after 6 months. Note that the further out the hazard over analysis time, the less reliable the estimate and thus any

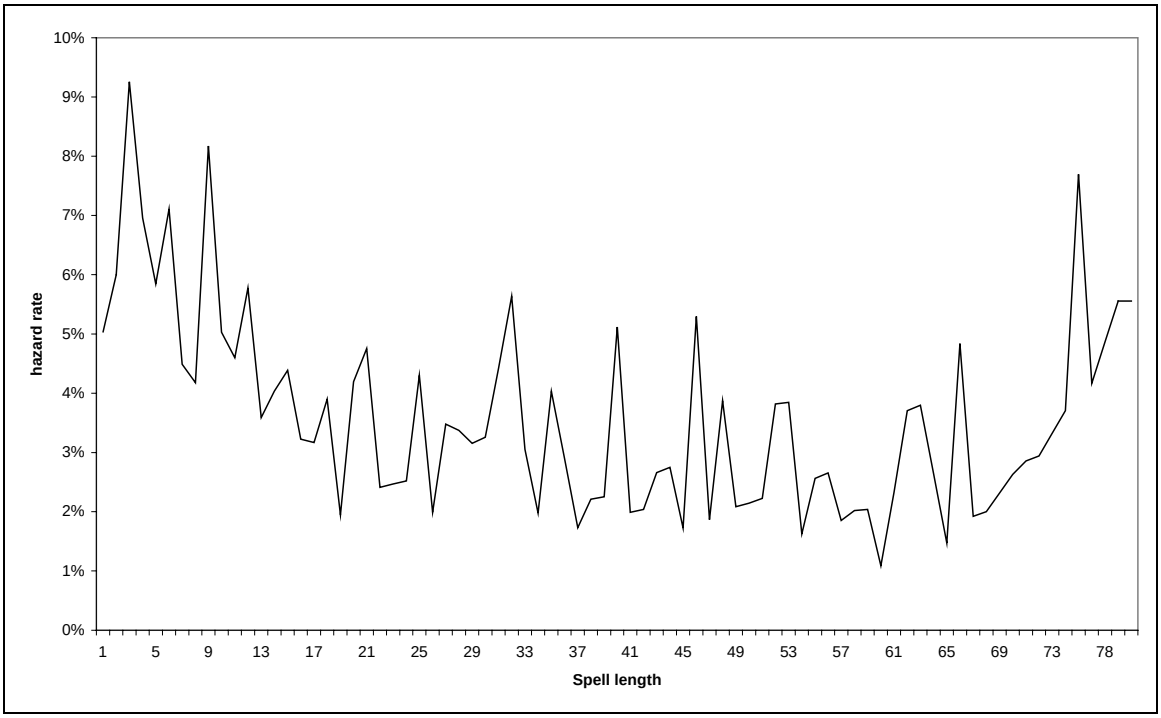
[10] While left censoring is a key issue in duration analysis, right censoring is less of an issue. This is because as long as right censoring is random (i.e. the probability that a spell of length t's exit will not be observed is random) then the spells enters into the denominator of time at-risk at each period, but because the spell is not observed to end, the spell never enters the numerator for spell end. Or in the likelihood function the survival term (rather than the density term) is included in the likelihood function.

conclusions about the peak after 75 periods (2.1 years) should be tentative. Another interesting point from figure 1 is the saw shape in the lines. That is we see peaks in the hazard rate after 3 periods (1 month), 6 periods (2 months) and so on. This most likely reflects the nature of the questions in HILDA and the fact that they were based on recall.

(i) *Can the calendar data from different interview dates be joined successfully?*

Ideally calendars reported at different interview dates (i.e. the wave 1 and the wave 2 interview dates) would always be consistent and there would be no large out-flow at the time of the join (the point where the calendars in wave 1 and wave 2 intersect). However, from figure 2 we can see that there is a considerable issue at the join, with a large fall in the number of people who remain unemployed and large increases in the in-flow and out-flow from unemployment. The number remaining unemployed falls from 500 per period to approximately 250 at the join, while the in-flow and out-flow to unemployment increase from less than 50 to nearly 250 per period at this time.

Figure 1: Raw hazard rate over time

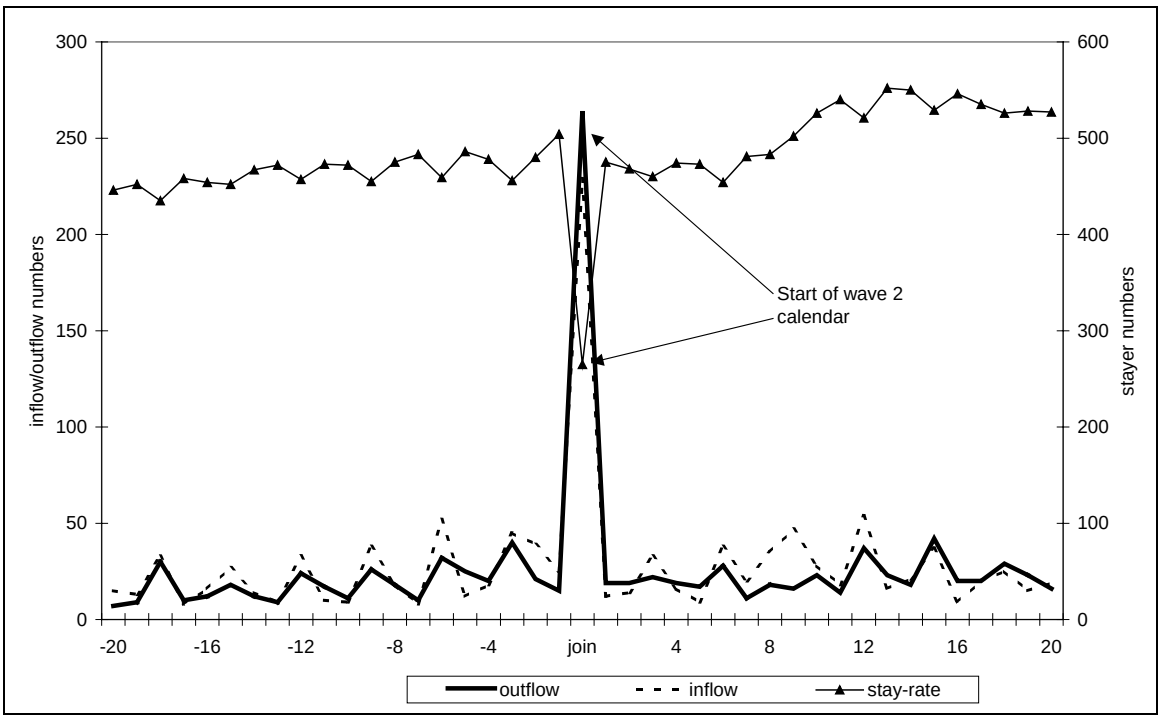


Approximately half of the unemployment periods recorded between July and November 2001 were reported inconsistently between waves 1 and 2 (there is an overlap

between the calendars of wave 1 and wave 2 of up to 5 months – 'the seam' - where we can check the consistency of answers between interview dates). I find that 27% of people gave answers between waves that were never consistent (where the categories are unemployment and not unemployment), 40% of people were consistent the entire time and the remaining 1/3rd were consistent for between 1 and 8 of the 9 periods examined.

Fortunately I can resolve the seam and join issues in a relatively straightforward manner. I apply the basic rule to the seam that the data reported closest to the interview data is used when there is some inconsistency (assuming that recall is better over shorter periods than longer periods). More importantly, where greater than 20% of observations are inconsistent between interview dates for times in the calendar, I treat the spells reported in wave 1 and wave 2 separately (because I consider that the quality of the join is too low). When I apply this rule the problem of the large in-flow and out-flow at the join is resolved.

Figure 2: Numbers entering, staying and exiting from unemployment



(ii) Are there issues with left censoring and left truncation?

Left censoring occurs when our first observation of the spell is sometime after it has begun and we do not know how long the spell has been going when it is observed.

The problem with left-censored unemployment spells is that their characteristics may be quite different to other unemployment spells. One possibility is to exclude the left censored data from analysis and under-take analysis with the inflows to unemployment. Another possibility is to use stock sample techniques to include the data (see Lancaster (1990)). However, stock sampling is not examined in most unemployment duration studies and is beyond the scope of this research.

Left truncated data are where the spell does not begin in the scope of the calendar, but we know the date it began. These data are included in the likelihood function here as the conditional probability of exit given the spell has lasted to length x (see Lancaster (1990)). In HILDA, the start date is available for spells that are in progress at the wave 1 interview date (15-18 months after the beginning of the reference period). If the spell began prior to the reference period and is continuous until the wave 1 interview date then the spell start date is available (left truncated data). Otherwise, if the spell began prior to the reference period and finished before the interview date then the start date is not available (left censored data). Because of this distribution between left censored and left truncated spells, the very longest spells in HILDA are likely to be primarily left truncated rather than left censored.

(iii) What variables are included in the analysis?

Explanatory variables available in HILDA related to wages include educational background, employment experience, location, country of birth, long term disability, parental employment status (which may affect opportunities and preferences for human capital) and previous unemployment experience. Explanatory variables available related to non-wage income (through intra-household income sharing and government transfers) and therefore reservation wages[11] include marital status, children, disability and location.

A number of our explanatory variables may be endogenous. Therefore it is important to use other information from the survey to backcast the variable to prior to the beginning of the spell. Back-casting is only done for spells that begin prior to the wave 1 interview date, otherwise the variable is as at the wave 1 interview date. Location is one example of a time varying endogenous variable. People living in high unemployment

[11] A reservation wage variable for the stock of unemployed is collected in HILDA at the interview date. However, these data are not used because the sample of reservation wages for the *inflow* to unemployment is too small and because the reservation wage variable is endogenous to the unemployment experience and thus it is not appropriate to use this reservation wage variable for estimation with the stock of unemployed (given their varying unemployment durations).

areas may be less likely to find work and the unemployed, who may have less income, may move to areas with cheaper housing but where there are fewer job prospects. HILDA has a question that asks how long has the respondent lived at their current address and using this information it is possible to confirm their location prior to the unemployment spell and partially overcome this endogeneity issue. Other variables treated in this fashion are number of children, marital status and employment and unemployment experience.[12]

It is not possible to back-cast the risk-aversion and lagged wage variables, which are potentially endogenous. In specifications that include these variables, I restrict the sample to include spells beginning after the wave 1 interview date (when the data are collected). A number of variables are primarily time invariant over the short period studied or not endogenous to unemployment and therefore they are recorded as at the wave 1 interview date. These variables are educational qualifications, parental employment status, sex, country of birth and disability.

I also include dummies to control for the time of interview. A dummy variable indicates if the spell began in wave 1, and two dummy variables indicate what part of the duration month the spell ended (monthly reporting dummies), to control for the saw shape in the raw hazard (see above).

5. Estimation of Full Model

The primary focus of this paper is what affects the probability that someone will move from unemployment to employment. I use the proportional hazards competing risks framework in estimation (see section III). I report here the results from the piecewise baseline hazard (with and without unobserved heterogeneity) and the Cox estimations (see table 2).[13] I present these three sets of results to show the robustness of the findings to alternative assumptions about the baseline hazard and because these estimations allow for more flexibility in the baseline hazard. The results from all estimations presented are consistent and in most cases the signs and levels of significance agree.

The coefficients are presented as hazard ratios. So a coefficient of, for example, 0.5 for a dummy variable is interpreted as lowering the exit rate from unemployment to employment by a half. For a continuous variable, a coefficient of 0.5 implies a unit

[12] Where a person moves location it is not possible to give their previous location, we simply have a missing location for this group. However, it is possible to back-cast children, marital status, and employment and unemployment experience.

[13] Carroll(2004) reports from the results from the parametric estimation with a number of alternative baseline hazards including the gamma and Weibull baseline hazards.

change in the variable is associated with a hazard rate 1/2 as large and an n unit change in the variable is associated with a hazard rate $(1/2)^n$ as large.

The preferred specification is the Cox partial likelihood estimator because, firstly, it does not rely on the restrictions of the baseline hazard. Secondly, the tests for violation of the proportional hazards assumption (see below) indicate that the wave 1 dummy and the monthly reporting dummies violate the proportional hazards assumption. The most convenient way to deal with this violation is with the Cox partial likelihood estimator stratified by wave1 and the monthly reporting dummies.[14]

The results described in this section are presented in table 2. We have 769 exits to employment and a total of 1400 spells overall for the continuous time models and 1396 spells for the discrete estimation (because the left truncated data are excluded). Looking at the means of the variables in the estimation,[15] we see that 54% of the spells are for males, 75% of the spells are for Australian born people, 16% are for people with university qualifications, 11% are experienced by people who have spent 25% of their work experience in unemployment (see column 4 of table 2).

(i) What impact do 'wage' variables have on exit rates to employment?

The first main result is that variables expected to increase wages are associated with higher exit rates from unemployment to employment (results from the preferred specification are presented in specification 1 in table 2). In particular, variables expected to increase wages through human capital and productivity enhancement or screening (employment experience and university qualifications) are associated with higher exit rates to employment.

We would expect productivity variables to increase the exit rates to employment because they increase the expected returns from working relative to not working. While they also increase the reservation wage (because they are associated with higher future earnings), Mortenson (1986) shows that variables expected to increase the wage are associated with shorter unemployment spells. Employment experience and university

[14] The saw shape presented in the earlier hazard curve in figure1 was held constant by dummy variables. It is not possible to stratify the parametric estimators by the monthly reporting dummies because this would result in discontinuities in the baseline hazard. It is possible to interact survival time with the monthly reporting dummies to overcome the violation, but initial interactions with the parametric estimator did not resolve the proportional hazards violation.

[15] These means were calculated for the continuous time sample, but there is little difference in the means for the discrete time sample (because of the large overlap between samples).

qualifications significantly increase the exit rate to employment at the 1% level of significance.

Having university qualifications increases the exit rate by 40% and shortens the median duration by approximately 40%.[16] One year of extra work experience increases the exit rate by 6%, while 10 years of extra work experience increases the exit rate by 72%. It should be noted that the addition of more work experience increases the exit rate at a decreasing rate (because the hazard rate of the square of employment experience is less than 1).

[16] The change in the median duration was calculated from the survivor function for the Cox estimation for a change from school qualifications to post-school qualifications. The effect of the variable on median duration may be sensitive to the value of the covariates. The covariate values are as follows: men, aged 40 years, with 5 years work experience, Australian born, with no disability, unmarried, in wave 1, with 25% unemployment experience (38.2% drop in median duration) and without 25% unemployment experience (39.9% drop in median duration).

Table 2: Affect of covariates on exit rates – Exits to employment - hazard ratio

	(1) Cox partial likelihood	(2) Piecewise w/out hetero.	(3) Piecewise with hetero.	Means for specification (1)
Male	0.825*	0.774*	0.659**	0.54
	[2.08]	[-2.53]	[-2.85]	
Non-English speaking COB	0.678*	0.547**	0.427**	0.14
	[2.02]	[-2.91]	[-2.59]	
English speaking COB	1.318	1.304	1.132	0.11
	[1.56]	[1.34]	[0.43]	
Length resident (Eng)	0.992	0.994	0.992	3.69
	[1.08]	[-0.74]	[-0.64]	
Length resident (non-Eng)	1.006	1.015	1.022	2.27
	[0.78]	[1.53]	[1.40]	
Non-university qual	1.007	0.971	0.912	0.44
	[0.11]	[-0.39]	[-0.85]	
University qual	1.398**	1.396**	1.842**	0.16
	[3.69]	[3.24]	[4.24]	
Father's unemploy status	1.278	1.163	1.112	0.13
	[1.76]	[0.92]	[0.47]	
Mother's employment status	0.941	0.926	0.971	0.44
	[0.86]	[-0.99]	[-0.26]	
Children	0.812**	0.753**	0.739**	1.19
	[4.17]	[-5.31]	[-4.30]	
Male*children	1.233**	1.326**	1.292**	0.60
	[3.55]	[4.54]	[2.98]	
Male*father	0.717	0.818	0.821	0.07
	[1.80]	[-0.93]	[-0.61]	
Marry	1.082	1.227*	1.365*	0.49
	[0.91]	[2.29]	[2.44]	
Disability	0.722**	0.718**	0.768	0.17
	[3.07]	[-3.09]	[-1.79]	
Employment experience	1.058**	1.059**	1.080*	12.97
	[3.20]	[2.89]	[2.47]	
Employment experience^2	0.999*	0.999	0.999	299.51
	[2.17]	[-1.93]	[-1.21]	
Lagged unemployment	0.731*	0.648**	0.436**	0.11
	[2.23]	[-3.15]	[-3.85]	
observations	1400	1396	1396	1400
failures (exits)	769	767	767	769
log likelihood	-3384.0	-2867.8	-2329.1	na
wald test	94.66	126.87	108.26	na
(prob> chi2)	(0.00)	(0.00)	(0.00)	

The base is women, without qualifications, Australian born, who are unmarried with no disability, with less than 25% of labour market experience in unemployment. In raw data there is a saw shape with systematically different exit rates for different parts of the month. Additional dummy variables and stratification in cox partial likelihood has been undertaken to control for this. Robust z statistics are reported in brackets. Issues remain with the non-proportionality of these variables in the parametric regressions (see discussion in text). Standard errors adjusted for clustering on xwaveid. Cox partial likelihood has been stratified to adjust for monthly reporting in calendar and for wave 1 dummy. Age, Age^2, regional unemployment rate and regional socio economic status are held constant. The discrete estimation was done with the baseline hazard specified as log time and left truncated data were excluded in estimation.

Being born in a non-English speaking country ("non-English speaking COB") is associated with lower exit rates from unemployment to employment. People born in non-English speaking countries may have problems with communication and culture (that may affect productivity or discrimination). Either way, we would expect this variable to lower wages and thereby lower exit rates from unemployment to employment. This variable significantly reduces the exit rate at the 5% level. The exit rates from unemployment to employment increase as people born in non-English speaking countries spend more time in Australia (although not significantly at the 5% level). Overall, being born in a non-English speaking country is associated with lower exit rates from unemployment initially, but as the immigrant spends more time in Australia the size of this effect diminishes.

The impact of disability on wages is likely to be both through productivity (if disability affects performance) and discrimination. We see that, as with the other wage variables, that disability takes the expected sign (having a disability lowers the exit rate). The coefficient on disability is significant at the 1% level of significance. Disability lowers the exit rate from unemployment to employment by 28% per period and increases the median duration by approximately 70%.[17]

(ii) What impact do non-wage income variables have on exit rates to employment?

The second main result is that variables associated with higher non-wage income (and hence we would expect higher reservation wages) are associated with lower exit rates from unemployment to employment. Two key sources of non-wage income are government transfers and intra-household income sharing. Two groups that in general receive higher Government transfers are people with children and people with disabilities.[18] Thus, we would expect these variables to be associated with lower exit rates (although these variables may also affect exit rates through productivity as well).

[17] A description of the covariates used for this calculation is given in footnote 16.

[18] To give an historical example of relative benefit levels, according to the Department of Family and Community Services, in the year to June 2001, the "Disability Support Pension" and the "Parenting Payment Single" benefit were A$402 per fortnight, while the "NewStart Allowance" (unemployment benefit) was A$322 per fortnight. Figures are for people aged 21-59 years living away from home. In addition, parents with children may be eligible for supplementary family payments and all benefit recipients may be eligible for rent assistance and other miscellaneous supplementary payments.

We see that having children for women and having a disability are both associated with lower exit rates from unemployment to employment.[19] This result is consistent with higher reservation wages being associated with lower exit rates from unemployment to employment. Interestingly, we see very little impact for men having children, perhaps suggesting that men either have less access to the non-wage income associated with children or that men with children tend to be more paid more (swamping the reservation wage effect).

We would expect that people who are married would have more access to non-wage income. There is no impact of marital status on the exit rate, although this may be related to non-observable characteristics of married people increasing the exit rate and swamping the reservation wage effect.

(iii) Does past unemployment experience affect exit rates to employment?

I now turn to look at lagged duration dependence, that is, does previous unemployment affect the exit rate in the current spell. Lagged duration dependence may occur because of discrimination against people with unemployment histories, and because of erosion of human capital and work habits, that all in turn lower wage offers and hence result in lower exit rates from unemployment.

The results from this estimation are consistent with unemployment experience before an individuals' current spell lowering the exit rates from unemployment to employment. This result is significant at the 5% level. The coefficient indicates that past unemployment experience lowers the hazard rate by 27%. This result holds employment experience constant. The combined effect of being out of work and unemployed is the combined effect of the "lagged unemployment" and the "employment experience" variables. Thus, there is an effect from not accumulating human capital on the job, as well as a separate negative impact from being unemployed and so the total impact would be larger than 27% (because it would need to take account of the loss in employment experience as well).

While we see evidence of lagged duration dependence, unobserved characteristics (such as unobserved motivation) may be correlated with both past unemployment experience and exits from unemployment in the current spell. Therefore, in this case it

[19] The lower exit rates from unemployment to employment for women with children and people with disabilities may also be related to lower wages offered to these groups (potentially because of discrimination).

may be that unobserved characteristics, rather then previous spells of unemployment, are associated with lower exit rates from unemployment. The key question for research then becomes did the characteristics develop over the previous unemployment spell, or did they exist prior to the beginning of that previous spell?

(iv) How robust are the results?

Firstly, when evaluating the robustness of the models in section V we see that the joint wald test of all coefficients equal to 0 is rejected at the 0.1% level of significance in all estimations in table 2. Now turning to the residual analysis, the Cox model fits the data well at short durations (see plot in Carroll (2004), which plots the generalised residuals against the integrated hazard).

The test of the proportional hazards assumption (using the Schoenfeld residuals) indicates the assumption is violated in the non-stratified estimation (also see Carroll(2004)). In particular, the reporting (to deal with the systematically different exit rates for different parts of the month) and wave 1 dummies (takes a value of 1 if the spell begins in wave 1) appear to exhibit significant non-proportionality. Stratification in Cox partial likelihood[20] estimation has been undertaken to control for this. When estimation is stratified by the reporting and wave 1 dummies, the proportional hazards assumption is no longer violated. Hence the Cox stratified estimation is the preferred estimation.

Investigation of the martingale residuals undertaken in Carroll (2004) did not point to any major issues with the estimation. Carroll (2004) investigates whether the male and female data should be pooled and finds that the coefficient on the interaction between male and children was statistically significant at the 1% level. However, it was also found that the joint test of significance of the other interactions between explanatory variables and male was not accepted at the 20% level of significance. Hence, in estimation the data are pooled but an interaction term between male and children is included.

(v) How do these results compare to earlier results from the literature?

The results presented in section V are consistent with the findings presented in Borland's (2000) review of the Australian literature. In general, exit rates to employment

[20] With stratified estimation the baseline hazard is allowed to vary for different values of the variable, but the coefficients on the other variables are the same across different values of the stratification variable.

are positively related to educational attainment and job experience and negatively related to non-wage income and reservation wages.[21]

One interesting and different result from the earlier literature is that when employment experience is added to the estimation, age becomes insignificant. This suggests that impact of this demographic characteristic on exit rates may be through the impact on employment experience, rather than through another transmission mechanism. Another interesting result from this paper is that past unemployment is associated with a lower hazard rate.

6. Inclusion of Lagged Wage and Risk Variables in Analysis

The results presented so far have provided confirmation that variables associated with higher wages and lower reservation wages are associated with increased exits from unemployment to employment. To get a more detailed understanding of the other factors that may be associated with exits from unemployment to employment I now report the results of two extended models that use additional variables from the wave 1 interview (see table 3). Because these variables are potentially endogenous to the unemployment experience I estimate these extended models on a sub-sample including only spells starting after the wave 1 interview date.

The first additional variable I include is a measure of financial risk. The risk-seeker variable is derived from question C6 in the self completion HILDA questionnaire. Where a person answers that they take substantial or above average financial risks then the variable takes a value 1, otherwise the variable takes a value of 0.[22] Six percent of the unemployed classify themselves as above average risk-seekers.

The second additional variable I include is wages earned in the year prior to the wave 1 interview date. The mean lagged annual wage is A$14,060 (this includes people who had no earned income over the previous year). I also include the percentage of the year worked to examine the wage effect separately from the % of year worked effect.

[21] Using a similar base category to Chapman and Smith (1992), I find a coefficient of 0.356 on educational qualifications and 0.470 on being Australian born (as opposed to being born in a non-English speaking country). Chapman and Smith (1992) find a coefficient of 0.272 on educational qualifications and 0.298 on being Australian born. The differences in size can be explained by the slightly different definitions used in the two papers and the positive coefficient on time spent in Australia for people born in non-English speaking countries in the current paper.

[22] While this risk variable may not be correlated with general risk-seeking behaviour (such as views regarding speeding), it is likely to be relevant to decisions about whether to decline earnings in the current period, for potentially higher earnings in the future.

Because risk and lagged wages are only available for analysis *after* the wave 1 interview, and because over half of the unemployment spells in the calendar begin prior to the wave 1 interview date, the sample size is reduced in these estimations. For this reason more caution should be used in interpreting the results. The results presented in this section are again presented in hazard ratio format. The number of cases is given towards the bottom of table 3.

The means of the variables for the extended models with the smaller sample are presented in column 5 of table 3. We see that the means of all the variables (university qualifications, marry, disability, employment experience and lagged unemployment) are similar between tables 2 and 3.

With the restricted sample a key question is whether the alternative sample leads to different results for the standard model. We see that the major difference between the standard models in tables 2 and specification 1 of table 3 is the coefficients are less precisely estimated as the standard errors are larger in table 3 (because the estimations in table 3 have a smaller sample size). In particular, the employment experience variable and the male variable become insignificant. However, overall the results from earlier estimations stand.

(i) Does a person's view of risk affect their exit rates to employment?

I now investigate whether or not risk plays a role in the rate of exit from unemployment to employment. Risk-seekers may be more likely to decline a *certain* wage offer in the current period, for the possibility of a higher *uncertain* wage offer in the future. We would expect risk seekers to have lower exit rates from unemployment to employment (holding other characteristics constant) because risk-seekers may be more likely to turn down wage offers. But, on the other hand, we would expect this group to have higher post unemployment wages.

Risk seekers are more likely to be male (68% compared to 52%), with greater income (37% earned more than $30,000 in the previous year compared to 25% of non risk seekers), university educated (24% compared to 13%) and with more employment experience (41% with more than 20 years employment experience compared to 33% of non risk-seekers).

Specification 2 from table 3 presents the results from the estimation when the risk–seeker variable is included in estimation. The coefficient on the risk-seeking variable is less than 1 and therefore indicates, as expected, that people who are prepared to take

financial risks are more likely to delay their exit from unemployment to employment to wait for a higher wage offer (hence they have a lower exit rate).[23] This result is not significant at the 5% level of significance, although the z statistic is 1.80 (however, when the lagged wage variable is added in the next section, we see that the risk variable becomes significant at the 5% level). This result suggests that on the supply-side, a key issue is how the unemployed view future wage offers (and in turn how they view risk and uncertainty).

To my knowledge there appears to have been very little, if any, consideration of risk aversion in the literature. However, given the possibility that perceptions of risk may influence the choice about whether to accept a certain wage offer in the current period or to decline and wait for an uncertain, but potentially higher, wage in a future period, this may be an interesting area of future research.

(ii) Does last year's wage explain the variation in exit rates?

I now investigate the role that previous wages may have on exit rates from unemployment (see specification 3 from table 3). This variable may be correlated with the wage distribution that people face in the current period. Both period's wage distributions are likely to be affected by the same unchanging unobservable characteristics, although some other characteristics may change somewhat between years. This gives us a further indication of the role that wage offers play in unemployment duration, beyond the indicators presented in section V. As noted above, I include the percentage of the year that individuals work, to control for the wage effect separately from the percentage of the year worked effect.

The coefficient on wages earned in the previous year is significant at the 1 percent level and indicates that the higher the wage earned in the previous year (holding the % of the year worked constant) the higher the exit rate from unemployment. If there is a high correlation between last year's wage offers and this year's wage offers, then this

[23] As with other variables in analysis, the risk seeker variable may be picking up unobserved characteristics that may be related to exit rates from unemployment. Notably, unemployed people who are financial risk-seekers may be more likely to have greater wealth (see risk-seeker characteristics above), and hence have higher reservation wages and lower exit rates. An alternative measure of risk-seeking behaviour is whether a person has ever smoked. This was included in estimation and while it was associated with a lower hazard rate, it was not found to be significant. There is a concern that while the smoking variable may be picking up some risk-seeking behaviour that it may also be related to a number of other factors that could be related to unemployment duration (such as parental income and education).

coefficient is consistent with higher wage offers being associated with higher exits from unemployment.

Interestingly, when we include this wage history variable, the coefficients on most of the wage-related variables (university qualifications, disability, employment experience) become insignificant at the 5% level and become smaller in sign (or switch sign). These correlations between wages in the previous year, exit rates to employment and the wage related variables provide further evidence that these variables affect exit rates through the wage offer distribution.

Table 3: Comparison of three different specifications (Cox) – Exits to employment - hazard ratios

	(1) Standard Model	(2) Risk Model	(3) Lag wage Model	Means for specification (2)
Male	0.835	0.884	0.794	0.52
	[1.36]	[0.89]	[1.54]	
Non-university qual	0.989	0.989	1.022	0.44
	[0.11]	[0.11]	[0.19]	
University qual	1.408*	1.409*	1.102	0.15
	[2.43]	[2.38]	[0.61]	
Children	0.749**	0.755**	0.748**	1.27
	[3.96]	[3.70]	[3.43]	
Male*children	1.282**	1.261**	1.291**	0.61
	[2.89]	[2.61]	[2.64]	
Marry	1.471**	1.459**	1.387*	0.43
	[3.28]	[3.13]	[2.49]	
Disability	0.636**	0.655**	0.797	0.21
	[2.98]	[2.73]	[1.33]	
Employment experience	1.040	1.030	0.992	12.84
	[1.56]	[1.19]	[0.29]	
Employment experience^2	0.999	0.999	1.000	295.16
	[1.29]	[0.95]	[0.05]	
Lagged unemployment	0.637**	0.651*	0.651*	0.15
	[2.77]	[2.57]	[2.26]	
Risk-seeking		0.685	0.576*	0.06
		[1.80]	[2.33]	
Lagged wage			1.011**	14.06
			[3.38]	
% of year worked			1.003	49.79
			[1.35]	
Controls for country of birth and parental employment status	yes	yes	yes	n.a.
Observations	959	894	793	894
failures (exits)	453	423	367	423
log likelihood	-2003.6	-1828.5	-1503.9	n.a.
Wald test	73.23	67.33	78.05	n.a.
(Prob > chi2)	(0.00)	(0.00)	(0.00)	

The base is women, without qualifications, Australian born, who are unmarried with no disability, with less than 25% of labour market experience in unemployment. Standard errors adjusted for clustering on xwaveid. Cox partial likelihood has been stratified to adjust for monthly reporting in calendar/wave 1 dummies. Estimation only undertaken on wave 2 data, because of endogeneity with explanatory variables. Age and Age^2 held constant. The means for male, non-university quals and country of birth not included here are similar to the estimates presented in table 4. Note the means for the wage variable are over a sample size of 791 rather than 892, because of the number of missing values on this covariate.

To check if the drop in significance is caused by near multi-collinearity between the lagged wage variable and the wage related variables I examined the correlation coefficients. There are high correlations between age and experience, and relatively high correlations between "lagged wage" and "employment experience" (but still a pairwise

correlation coefficient of less than 0.5). Overall, the pairwise correlation coefficients do not indicate a problem with near multi-collinearity.

7. The Underlying Baseline Hazard

A usual finding when negative duration dependence is observed is that it is not possible to distinguish whether longer spells result in lower exit rates, or whether there is unobserved heterogeneity leading to 'low exiters' remaining in unemployment for longer. However, the shape of the baseline hazard is important, because falling exit rates by duration (holding characteristics constant) indicates that unemployment has a scarring effect, which suggests adjustment may come with medium-term costs. One additional point to note is that steady exit rates with duration may occur because of factors offsetting unemployment scarring effects.[24]

(i) Do hazard rates from unemployment to employment decline with duration?

Our data show that the exit rates from unemployment to employment do not initially appear to decline with duration (for the first four months), but then hazard rates appear to decline relatively sharply.

Figure 3 shows the baseline hazard curves for the raw data, the Weibull, the piecewise constant (and the confidence bands for the piecewise constant hazard). The Weibull has been the more traditional of the measures to examine the baseline hazard in Australia, however, because it summarises duration dependence into one parameter, interest in this measure has waned. The preferred measure of the time pattern of hazard rates is the piecewise constant hazard, which shows an initially flat hazard curve before declining relatively sharply after four months. However, it should be noted that the only statistically significant time parameter is on the one year+ duration effect.

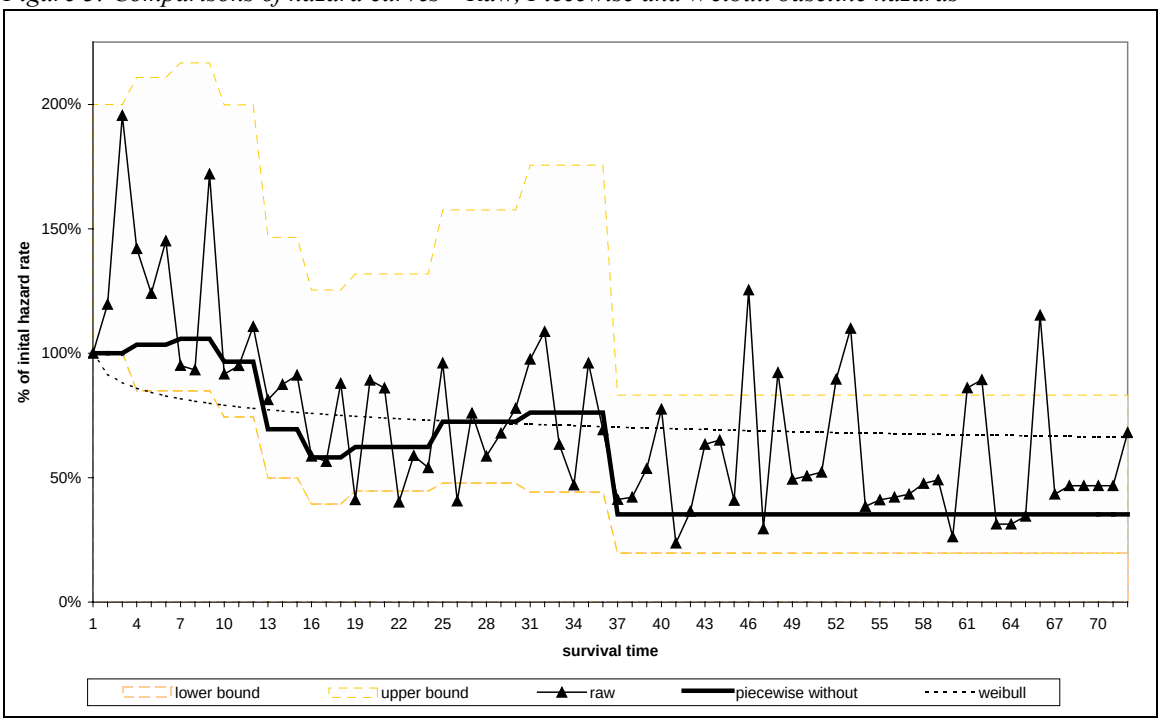
The sharp decline in the hazard rate after four months may reflect unobserved heterogeneity or state dependence effects.[25] Thus, as stated above, there is interest in understanding whether the heterogeneity developed prior to the start of the spell

[24] Machin and Manning (1999) highlight that the presence of unemployment benefits that decline with duration and active labour market policies targeted at long-term unemployed (amongst other factors) may lead to rising exit rates with duration. Thus, even minimal observed negative duration dependence may be associated with scarring effects of unemployment (because scarring is off-set by institutional factors).

[25] One method to control for the unobserved heterogeneity is to integrate it out by restricting the distribution of the unobserved heterogeneity to be of a certain shape. Some studies have concluded that restricting the unobserved heterogeneity distribution is very sensitive to the restrictions used and is not necessarily better than excluding the term where a flexible baseline hazard is used (see Narendranathan and Stewart (1993) and Boheim and Taylor (2000)).

(unobserved heterogeneity), or during the spell itself (state dependence). If there is state dependence, the policy response (if any) will be determined by the drivers of the state dependence (human capital depreciation, discrimination by employers, or discouragement).

Figure 3: Comparisons of hazard curves – Raw, Piecewise and Weibull baseline hazards[26]



The non-monotonic hazard may occur because people are more likely to forget short spells of unemployment (<1 month). If the issue is a reporting one then the 'true' distribution may be monotonic (and we could potentially use a Weibull baseline hazard). However, estimation should then take this into account and investigation has shown that the estimates of the Weibull shape parameter are sensitive to the inclusion of the first two periods data.[27] There may also be genuine reasons to believe that we may observe flat hazard rates at the beginning of a spell. One example is that job search may require some learning (for example CV creation, identification of employers to approach) and thus people's job search may initially improve (leading to an initially rising hazard rate), before any negative effects of long-term unemployment occur.

[26] To control for the fact that people are more likely to report unemployment for a full month, the time bands used for analysis are: 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7-8 months, 9-10 months, 11-12 months, 12< months.

[27] Where we exclude the first 2 periods data the implied fall in the exit rate between 1 month and 12 months is 48% compared to 19% when the first 2 periods of data are included.

(ii) How much of the decline in the hazard is explained by the covariates?

The previous sub-section showed that in general there appeared to be a downward slope in the baseline hazard after four months of elapsed duration, but it is not possible to know whether this shape is because of state dependence or unobserved heterogeneity. A natural next question is how much of the falling hazard rates did we explain with the variables included in the estimation with a piecewise constant baseline hazard.

Given that we have included demographic information, unemployment history, educational qualifications, whether Australian born and a variety of family variables, we may think that we would explain a significant part of the variation. On the other hand it may be that the unobserved characteristics (such as motivation, natural ability) are important drivers of the wage distribution and of exit probabilities.

I find that observed characteristics explain less than 15% of the decline in the hazard rate observed between 1 month and 12 months in the piecewise constant model (see table 4). Table 4 shows how the baseline hazard varies with duration. Using the duration parameters from the piecewise baseline hazard without covariates the hazard declines by 46% over the first six months and a further 27 percentage points over the second six months, while when covariates are added the hazard declines by 42% over the first six months and a further 23 percentage points the second six months.[28] The large component of the decline in the hazard rate that is not explained by observable characteristics suggests either that unobservable characteristics are particularly important, or that some scarring is occurring as well as some unobserved heterogeneity.

Table 4: Investigating the piecewise coefficients – how do hazard rates vary with duration?

	Simple piecewise	Full piecewise
1 month (3 periods)	100%	100%
3 months (9 periods)	99%	106%
6 months (18 periods)	54%	58%
12 months (36 periods)	27%	35%
% of 12 months+ effect explained by explanatory variables		11.3%

The simple model excludes all explanatory variables. The full model includes all explanatory variables included in specification 2 of table 2.

[28] Note that the value of the hazard will vary with the values of the covariates, but the relative decline in the hazard will be constant across estimates.

This result highlights that, even holding constant employment experience, qualifications, previous unemployment experience and other factors, there is a large difference in exit rates between people at longer durations than those at shorter durations. This also highlights how large the unobserved component would have to be for there to be no real fall in the hazard rate with duration.

8. Conclusion

This paper found that variables associated with increased wages are associated with higher exit rates from unemployment to employment. These wage variables include employment experience, educational qualifications. I also found that variables expected to lower wages (through impacts on productivity or discrimination) are associated with lower exit rates (non-English speaking COB, disability, lagged unemployment). Variables that may be associated with increased non-wage income (children, disability), and thus higher reservation wages, are associated with lower exit rates from unemployment. Reassuringly, results were robust to a variety of different estimation methods and specifications.

I found that people who are risk-seekers were more likely to experience longer spells of unemployment. This finding is consistent with risk-seekers being more likely to decline a certain current wage for a potentially higher future uncertain wage. This is an interesting result that deserves further investigation, because the key decision a job seeker may face is the choice about whether to accept a *certain* wage offer in the current period or to decline and wait for an inherently *uncertain* wage in a future period.

I found that the piecewise constant hazard indicated that the shape of the hazard curve may be flat at shorter durations, before declining after four months, as illustrated by the piecewise constant baseline hazard. Finally, while the duration dependence observed may be the result of unobserved heterogeneity or state dependence, the paper showed that less than one fifth of the decline in the hazard observed is explained by the inclusion of explanatory variables. This suggests that either unobserved characteristics are potentially important or that some scarring is occurring.

References

Arulampalam, W. (2001), "Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages", *The Economic Journal*, Volume 111(475), 585.

Arulampalam, W., Booth, A. and M. Taylor (2000), "Unemployment persistence", *Oxford Economic Papers 52.*

Arulampalam, W. and M. Stewart (1995), "The Determinants of Individual Unemployment Durations in an era of High Unemployment", *Economic Journal*, 105 (321-332).

Barrett, G. (2000), "The effect of educational attainment on welfare dependence: Evidence from Canada", *Journal of Public Economics* 77.

Böheim, R. and M. Taylor (2000), "Unemployment duration and exit states in Britain", University of Essex working paper.

Borland, J. (2000), "Disaggregated Models of Unemployment in Australia", Melbourne Institute Working Paper No. 16/00.

Brooks, C. and P. Volker (1986), "The Probability of Leaving Unemployment: The Influence of Duration, Destination and Demographics", *Economic Record*, September 1986, vol. 62.

Carroll, N. (2004), "Explaining Unemployment Duration in Australia: Data, methods and additional results", unpublished mimeo (available at http://econrsss.anu.edu.au/Students/carroll/workpapers_nc.htm).

Chapman, B. and P. Smith (1992), "Predicting the Long-Term Unemployed: A Primer for the Commonwealth Employment Service", in Gregory, R.G. and T. Karmel (eds), *Youth in the Eighties – Papers from the Australian Longitudinal Survey Research Project*.

Clark, A. and A. Oswald (1994), "Unhappiness and unemployment", *The Economic Journal*, 104(424).

Dockery, M. (2003), "Looking Inside the Unemployment Spell", Paper presented to the National Conference on the Household, Income and Labour Dynamics in Australia Survey (HILDA).

Heckman, J. and B. Singer (1984), "A method for minimising the impact of distributional assumptions in econometric models for duration data", *Econometrica*, 52(2).

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

Machin, S. and A. Manning (1999), "The causes and consequences of long-term unemployment in Europe", in O. Ashenfelter and D. Card (eds) *Handbook of Labor Economics* Volume 3C (Amsterdam, North Holland), 3085-3139.

Mortensen, D. (1986), "Job Search and Labour Market Analysis", in *Handbook of Labour Economics*, Vol.2.

Narendranathan, W. and M. Stewart (1993), "Modelling the probability of leaving unemployment: competing risks model with flexible baseline hazards", *Applied Statistic*s, 42(1).

Ruhm, C. (1991), "Are Workers Permanently Scarred by Job Displacements?", *American Economic Review*, 81(1).

Winkleman, L. and R. Winkleman (1998), "Why are the Unemployed so Unhappy: Evidence from Panel Data", *Economica*, 65(257).

Appendix I: Variable descriptions

Variable	Description[29]
Age	If spell begun prior to calendar then age at wave 1 interview date minus 2, if spell begun prior to wave 1 interview date then age at wave 1 interview date minus 1, If spell begun after wave 1 interview date then age at wave 1 interview date minus 1.
Wave 1	=1 if spell begun in wave 1 calendar, =0 otherwise.
Male	=0 for female, =1 for male
Regional unemployment	regional unemployment rate (where region is capital city versus rest of state).
SES decile	Regional SEIFA 2001 Decile of Index of relative socio-economic disadvantage. Refers to the region the person was immediately prior to beginning spell. Where the person moved between unemployment spell begin date and interview date the data will be missing. Regional disaggregation is finer than that used for regional unemployment.
Non-English speaking COB	=1 if born in a non-English speaking country, =0 otherwise
Length resident (non-Eng)	Length of time spent in Australia for people born in non English speaking countries.
English speaking COB	=1 if born in an English speaking country (ex. Australia), =0 otherwise
Length resident (Eng)	Length of time spent in Australia for people born in English speaking countries (excluding Australia).
Non-university qual	=0 if no qualifications recorded, =1 if non university qualifications recorded
University qual	=0 if no university qualifications, =1 if university qualifications recorded
Father's unemploy status	=0 if father not unemployed for a total of 6 months growing up =1 if father not unemployed for a total of 6 months growing up
Mother's employment status	=0 if mother not in employment when respondent aged 14 years =1 if mother in employment when respondent aged 14 years
Children	Number of children as at the beginning of unemployment spell. Note: children's age at the interview date was used in calculation
Marry	Marital status as at the beginning of the unemployment spell (either married/ de facto married, or not married)
Disability	=0 if no long term disability recorded at wave 1 interview =1 if long term disability recorded at wave 1 interview
Employment experience	Number of years in employment prior to unemployment spell
Employment experience^2	Number of years in employment squared prior to unemployment spell
Lagged unemployment	=1 if person spent more than 25% of time since school in unemployment =0 otherwise
Age^2	Age squared
Male*children	Interaction between male and number of children
Male*father	Interaction between male and father's unemployment status
mark1	Observation recorded in first third of month
mark2	Observation recorded in second third of month
Lagged wage	Respondents wage (in thousands) in year prior to 1st interview date
% past year in job	% of the past year that the respondent has been in paid employment (asked at wave 1 interview date)
Risk-seeking	=1 if respondent takes substantial or above average financial risks =0 otherwise (asked at wave 1 interview date)

[29] As at wave 1 interview date unless otherwise specified.