

TOWARDS A FORMAL PROFILING MODEL TO FOSTER ACTIVE LABOUR MARKET POLICIES IN NEW ZEALAND

James Obben¹

Department of Applied and International Economics
Massey University
Palmerston North
New Zealand

ABSTRACT

Following the unemployment hysteresis of the 1980s, discussions of methods for reducing the natural rate of unemployment tend to focus on long-term unemployment (LTU). A broad consensus exists among economists for efforts to move towards active labour market policies or ALMPs that increase employment opportunities. The premium is on accurate and early identification of jobseekers at risk of LTU (an activity known as profiling) so that action on referral to an appropriate ALMP can be taken as early as possible in the unemployment spell. Evidence shows that statistical-model based profiling is more accurate than profiling based on other methods. To advance a statistical profiling model for New Zealand, this paper has attempted to fit a logit model to more than 2.2 million completed unemployment spells. Despite the prevailing data limitations, the estimated coefficients have the expected signs and form the most plausible basis yet for estimating the probability an identified jobseeker has of suffering LTU.

JEL Classification: J64

1. INTRODUCTION

Much as it is decried, some unemployment is necessary for any market economy to function properly (Layard *et al.* 1991, p. xiii). Thus a lot of attention in the literature has been directed at measuring, describing and explaining unemployment, especially its equilibrium level – the level at which inflation stabilises.² Ever since the introduction of the natural rate hypothesis (NRH) and the expectations-augmented Phillips curve in seminal papers by Phelps (1967) and Friedman (1968), the concept of the ‘natural’ rate of unemployment or the NAIRU (the non-accelerating inflation rate of unemployment) has dominated the discussions of equilibrium unemployment.³ In its simplest version, the NRH posits that in the absence of supply shocks unemployment will deviate from the

¹ I wish to thank Andy Reynolds who made it possible for me to obtain the data. Ted Drawneek provided invaluable computing assistance in editing the data for analysis. Comments on an earlier draft by Jim Alvey and S. Shakur were helpful and are appreciated. Any remaining weaknesses or errors are my responsibility

² Nickell (1990) surveys the vast literature on unemployment theories.

³ See the volume edited by Cross (1995).

natural or equilibrium rate only if expectations about inflation are wrong (i.e., inflation expectations error has a negative effect on unemployment levels). The natural rate or NAIRU is an attractor (but might also be an “attractee”) of the actual rate of unemployment. The main policy implication is that if unemployment is pushed below the NAIRU, inflation pressure builds up; and if unemployment stays above the NAIRU, disinflation ensues. It is only at the NAIRU that there is no tendency for inflation to either increase or decrease. The problem is, estimation of the NAIRU is quite imprecise (Staiger *et al.*, 1997) and there is controversy as to its practical usefulness (Galbraith, 1997). That notwithstanding, estimates of the time-varying NAIRU (Gordon, 1997) are watched with keen interest by country administrations and central banks (Stiglitz, 1997 and Richardson *et al.*, 2001). If estimates of the NAIRU are lower than the prevailing unemployment rate, it may be inferred that expansionary policies may be pursued without igniting inflation; if the natural rate rises and it is used as a serious “speed limit” to stop the reduction of unemployment in an anti-inflation policy, then relatively high unemployment rates may be interpreted as being normal or natural.

Unemployment in all the major industrial economies rose steadily from the mid-1970’s, stayed high in the 1980’s and 1990’s in most European countries (even when inflation stabilised) but fell in the US, Japan and a few other countries (Martin, 1994). Since the NRH argues that the unemployment rate over long periods cannot move too far from the natural rate, the persistently high observed unemployment meant the natural rate rose. To help explain this phenomenon, some economists have resorted to the theory of “hysteresis” – the notion that long-lasting recessions leave relatively permanent effects on the time path of the economy.

There are a number of ways in which unemployment hysteresis might occur. In the view of Layard *et al.* (1991) a key to understanding this is the emergence of long-term unemployment⁴ (LTU) that accompanied the high unemployment in Europe. A deep recession displacing a lot of workers for extended periods means the proportion of long-term unemployed workers among the unemployed increases. If the unemployment benefit scheme is quite generous, there is little incentive for jobseekers to accept low-paying jobs and might prefer remaining unemployed for a long time. If the long-term unemployed become demoralised and search less effectively or lose their skills, employers may not find them employable and may in fact discriminate against hiring the long-term unemployed applicants (Nickell, 1990, p. 420). If wage setting is largely through union bargaining, then outsiders cannot exert any downward pressure on wages and thereby bring about an increase in employment: insiders will pursue high real wages for themselves leaving outsiders to remain unemployed. Unemployment (which is the same as excess supply of labour) exerts downward pressure on wages. However, the higher the proportion of long-term unemployed, the less effective is unemployment in holding down wages and helping to lift employment. Thus discussions of methods for reducing the natural rate tend to focus on LTU.

In focussing on LTU to reduce the natural rate, there is a broad consensus among economists for efforts to move away from the “passive” help given through unemployment benefits to “active labour market policies” or ALMPs that increase employment opportunities. Three mechanisms are identified here: (i) job-search assistance or placement and counselling services, (ii) education and training, and (iii) direct job creation and recruitment subsidies (Layard *et al.*, 1991; Katz, 1994).

⁴ In New Zealand, long-term unemployment is defined as continuous unemployment spell lasting 27 weeks or more in contradistinction to the OECD definition that uses 52 weeks.

The details and relative merits of these are discussed in the references above. If these policies are targeted successfully and in a cost-effective manner at the long-term unemployed or those jobseekers at risk of being long-term unemployed, it should be possible then to remove more people from the unemployment register sooner rather than later. Certainly the premium is on accurate and early identification of jobseekers at risk of LTU (an activity known as “profiling”) so that action on referral to an appropriate ALMP can be taken as early as possible in the unemployment spell.

Profiling may be based on either (a) statistical models estimated to determine the probability of an unemployment insurance claimant exhausting their benefit [as in the US] or to identify those jobseekers who have the highest risk of LTU [as in Australia], or (b) “characteristics screening” – sorting jobseekers according to whether or not they belong to groups with known unemployment handicaps, or (c) the judgement of the public-employment-service staff. Even though the predictive power of statistical models is relatively low, evidence from Australia and the United States show that model-based profiling is more accurate than profiling based on judgement of public-employment-service staff and on characteristics screening (OECD Proceedings, 1998, p.16).

New Zealand, along with the other OECD countries, experienced increases in unemployment and LTU in the 1980’s and 1990’s (see Table 1). Indeed, during the early 1990’s New Zealand policy makers (Employment Task Force, 1994) identified the growth of LTU as the single most worrying feature of the country’s recent economic and social performance (see the incidence of LTU in New Zealand during the 1987-99 period summarised in Tables 2 and 3). In response to these developments some New Zealand economists directed research effort towards analysing the New Zealand Employment Service (NZES) dataset. Using survival analysis techniques and censored data (i.e., a mixture of completed and uncompleted unemployment spells), the objectives of those studies have been to determine the extent to which the length of unemployment duration can be predicted from the characteristics of the unemployed individuals (e.g., Watson *et al.*, 1997) and to determine the impact that certain characteristics and unemployment duration have on the probability of someone exiting the unemployment register (e.g., Gardiner, 1995). Unlike in Australia where formal profiling using a statistical model to identify at-risk jobseekers for early intervention was initiated in 1994, in New Zealand assistance is provided to registered individuals when they are clearly identified as long-term unemployed. Some concerns about the timing of assistance and early identification of at-risk jobseekers had been raised by the Prime Ministerial Task Force on Unemployment (1994) but to date no formal profiling is done.

To advance a statistical profiling model for New Zealand, this paper extends a previous work (Obben *et al.*, 2002) that estimated a logit model based on a “smaller” random sample of 100,000 completed unemployment spells (i.e., uncensored data) from the NZES database traversing the 1988-97 period to cover more than 2.2 million observations available for the period. Earlier studies modelling unemployment duration in New Zealand had used censored data but, as Salant (1977) has noted, parameters based on incomplete spells suffer from interruption bias and length bias (or the censoring problem). Utilising completed spells or uncensored data thus skirts the objection raised by Salant and also allows all the cases to be neatly categorised either as being long-term unemployed or otherwise, thus making it possible to apply binary choice models as an alternative approach. Barron and Mellow (1981) and Heikensten (1984) are two examples of studies that used the binary choice approach to estimate re-employment probability.

Table 1: Incidence of Long-Term Unemployment (as Percentage of Total Unemployment) in Surveyed Selected OECD Countries 1988-99

Country		1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Australia	≥6mths	47.4	40.6	41.0	49.6	58.7	57.1	41.0	51.3	48.5	51.4	52.2	48.4
	≥12mths	28.4	23.0	21.6	24.9	34.5	36.5	36.4	30.8	28.4	30.7	33.6	29.4
Austria	≥6mths								42.8	42.5	47.7	43.3	47.6
	≥12mths								25.6	25.6	28.7	29.2	31.7
Belgium	≥6mths	90.0	87.5	81.4	77.9	74.7	70.4	75.2	77.7	77.3	77.2	76.3	73.5
	≥12mths	77.5	76.3	68.7	62.9	59.0	52.9	58.3	62.4	61.3	60.5	61.7	60.5
Canada	≥6mths	20.7	20.8	20.2	22.5	26.1	28.6	28.4	27.8	29.3	26.9	24.1	21.4
	≥12mths	7.4	6.8	7.2	6.5	9.4	11.4	12.5	14.1	16.7	16.1	13.7	11.6
Czech Rep	≥6mths							40.9	52.5	52.3	53.0	54.6	61.9
	≥12mths							21.6	30.6	31.3	30.5	31.2	37.1
Denmark	≥6mths	51.7	50.8	53.2	54.4	49.9	45.5	54.0	46.6	44.4	45.7	41.4	38.5
	≥12mths	28.7	25.9	29.9	31.9	26.9	25.2	32.1	27.9	26.5	27.2	26.9	20.5
Finland	≥6mths	...	23.6	32.6	32.6	...	52.8	52.8	56.6	55.5	48.6	42.2	46.4
	≥12mths	...	6.9	9.2	9.2	...	30.6	30.6	37.6	34.5	29.8	27.5	29.6
France	≥6mths	64.6	63.7	55.5	58.0	58.1	58.2	61.7	64.0	61.5	63.7	64.2	55.5
	≥12mths	44.8	43.9	38.0	37.2	36.1	34.2	38.3	42.3	39.5	41.2	44.1	40.3
Germany	≥6mths	65.0	66.7	64.7	54.1	55.4	60.1	63.8	65.9	65.3	68.5	69.6	67.2
	≥12mths	46.7	49.0	46.8	31.5	33.5	40.3	44.3	48.7	47.8	50.1	52.6	51.7
Greece	≥6mths	71.3	73.5	71.9	71.5	70.4	71.0	72.8	72.7	74.7	76.5	74.8	...
	≥12mths	48.1	52.4	49.8	47.6	49.7	50.9	50.5	51.4	56.7	55.7	54.9	...
Hungary	≥6mths							62.6	73.0	75.2	73.5	71.0	70.4
	≥12mths							41.3	50.6	54.4	51.3	49.8	49.5
Iceland	≥6mths			13.6				31.4	33.3	31.2	27.0	22.9	20.2
	≥12mths			6.7				14.3	16.8	19.8	16.3	16.1	11.7
Ireland	≥6mths	82.2	82.4	81.0	77.6	77.4	76.9	80.7	77.9	75.7	73.6
	≥12mths	66.0	67.3	66.0	61.6	58.9	59.1	64.3	61.4	59.5	57.0
Italy	≥6mths	85.9	58.7	85.2	84.3	69.8	76.5	79.5	80.2	80.8	81.8	77.3	77.2
	≥12mths	69.0	70.4	69.8	68.1	58.2	57.7	61.5	63.6	65.6	66.3	59.6	61.4
Japan	≥6mths	40.5	37.3	39.0	38.1	36.2	34.4	36.1	37.7	40.4	41.3	39.3	44.5
	≥12mths	20.2	18.7	19.1	17.9	15.9	17.2	17.5	18.1	20.2	21.8	20.9	22.4
Korea	≥6mths			13.9				20.6	16.9	16.0	15.8	14.7	18.6
	≥12mths			2.6				5.4	3.4	3.9	2.6	1.6	3.8
Luxembourg	≥6mths	57.7	57.1	66.7	47.4	36.7	62.2	54.7	49.5	44.6	61.1	55.2	53.8
	≥12mths	42.3	38.1	42.9	26.3	13.3	32.4	29.6	23.8	27.6	34.6	31.3	32.3
Mexico	≥6mths								7.9	9.8	6.9	3.3	6.8
	≥12mths								1.5	2.2	1.8	0.9	1.7
Netherlands	≥6mths	65.6	66.1	63.6	60.3	76.8	79.1	77.5	80.4	81.8	80.4	83.5	80.7
	≥12mths	50.0	49.9	49.3	46.1	43.9	52.3	49.4	46.8	50.0	49.1	47.9	43.5
New Zealand	≥6mths	29.4	34.5	39.5	39.0	53.2	52.5	50.0	43.3	36.6	36.4	37.9	39.0
	≥12mths	11.3	14.7	20.9	21.2	31.9	33.2	32.2	25.5	20.7	19.4	19.4	20.8
Norway	≥6mths	15.9	29.5	40.8	39.1	41.1	45.6	43.8	39.1	31.1	26.1	20.5	16.2
	≥12mths	6.3	11.6	20.4	20.2	23.5	27.2	28.9	24.1	16.0	12.0	8.2	6.8
Poland	≥6mths			62.8				65.1	63.1	62.2	60.4
	≥12mths			39.0				40.3	40.0	38.0	37.4
Portugal	≥6mths	67.0	66.6	62.4	58.4	37.5	45.2	57.2	65.1	66.7	66.7	64.5	63.8
	≥12mths	51.2	48.3	44.8	38.7	31.0	43.4	43.4	50.9	53.1	55.6	44.7	41.2
Spain	≥6mths	75.0	72.7	70.2	68.4	66.1	69.6	73.4	72.8	72.2	71.8	70.4	67.9
	≥12mths	61.5	58.5	54.0	51.1	47.4	50.1	56.1	56.9	55.7	55.5	54.1	51.3
Sweden	≥6mths	21.2	18.4	22.2	17.6	25.9	32.0	38.7	45.6	48.4	50.8	49.2	...
	≥12mths	8.2	6.5	12.1	4.2	8.3	10.9	17.3	27.8	30.1	33.4	33.5	...
Switzerland	≥6mths			26.2	27.5	37.8	47.5	49.3	50.4	52.1	49.4	48.9	61.0
	≥12mths			16.4	17.0	19.6	20.2	27.6	33.6	25.0	28.5	34.8	39.8
Turkey	≥6mths			72.6				68.5	60.3	65.9	62.7	60.3	60.6
	≥12mths			47.0				45.4	36.3	43.5	41.6	40.0	33.7
United Kingdom	≥6mths	61.5	57.2	50.3	47.2	57.3	62.9	63.4	60.8	58.1	54.8	47.3	45.7
	≥12mths	44.7	40.8	34.4	28.8	35.4	42.5	45.4	43.6	39.8	38.6	32.7	29.8
United States	≥6mths	12.1	9.9	10.0	13.0	20.6	20.4	20.3	17.3	17.5	15.9	14.1	12.3
	≥12mths	7.4	5.7	5.5	6.3	11.2	11.7	12.2	9.7	9.5	8.7	8.0	6.8
Total OECD	≥6mths			44.6				52.4	49.8	50.3	50.8	48.3	46.2
	≥12mths			30.9				35.2	33.9	34.2	34.9	33.1	31.2

Source: OECD Employment Outlook, various issues.

Table 2: Percentage Share of Long-term Unemployment¹ in the Total Unemployment of Various Demographic Groups in New Zealand, 1987-1999

Group	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Age Group													
15-19	17.4	19.6	21.0	24.2	26.2	34.8	37.4	34.1	28.2	23.7	20.2	20.2	23.1
20-24	14.7	21.2	31.2	30.6	32.2	42.0	46.9	45.0	39.0	27.4	30.2	27.8	29.4
25-34	18.2	20.6	31.0	33.1	37.3	44.3	50.4	49.0	43.3	35.7	31.4	32.0	31.0
35-44	21.4	25.3	30.2	35.1	41.2	46.9	54.7	52.7	48.8	43.5	37.3	40.5	42.4
45-54	30.0	30.4	32.6	45.5	45.1	52.8	60.0	60.8	58.8	50.4	48.3	46.6	45.9
+55	...	39.3	41.3	43.4	46.9	56.9	65.9	65.8	62.7	60.0	47.1	47.9	52.2
Gender													
Male	27.1	31.5	37.7	42.7	43.8	52.1	58.2	57.0	53.8	44.7	40.9	39.4	42.3
Female	22.2	23.7	28.7	30.4	33.9	40.1	48.3	45.8	40.7	35.0	31.6	31.9	33.3
Ethnic Group													
European													
Maori	17.6	20.8	27.6	31.9	31.8	39.4	46.6	45.3	40.1	33.7	30.7	29.7	32.0
Pacific Isld	23.4	26.8	32.7	31.7	43.7	51.9	53.5	52.8	48.3	37.0	36.8	36.7	40.2
Other	32.4	41.3	46.2	57.4	68.6	58.6	56.7	50.5	43.3	39.8	39.1
	28.2	31.1	34.3	41.2	45.1	52.2	41.5	39.7	33.3	40.0	40.2
Education Quals.													
No quals.													
Sch. quals.	22.7	28.0	33.4	39.5	45.1	54.1	58.4	56.7	52.9	44.5	41.9	43.4	42.7
Post-sch.	13.6	16.1	24.1	26.7	26.8	34.5	41.9	39.1	33.9	28.4	23.0	23.8	29.4
Other quals	22.0	24.1	31.5	36.5	41.9	47.3	58.3	58.0	52.7	42.6	36.6	35.8	41.1
	13.6	16.3	23.2	18.5	25.7	35.1	40.9	39.9	35.3	28.9	29.3	26.8	28.3
Occupation Sought²													
MP&T ³						36.6	44.1	43.5	36.4	34.1	28.1	29.8	32.4
Clerks						39.1	50.6	46.9	45.4	38.2	30.8	28.6	33.1
S&SW ⁴						39.6	44.9	41.9	34.9	27.6	29.4	29.8	28.8
A&FW ⁵						46.9	50.0	49.5	44.9	44.0	41.3	40.7	41.8
Trades ⁶						47.3	58.6	54.8	60.3	41.8	37.5	33.3	38.9
P-MO&A ⁷						52.4	59.4	55.0	53.4	40.6	42.9	43.1	43.1
Elem. Occ ⁸						57.5	60.3	57.3	51.7	48.8	43.5	40.9	44.0
Not specifd						41.5	43.8	42.5	38.9	33.8	31.1	31.8	34.5
Overall	19.1	22.3	29.0	32.7	35.7	43.9	49.9	48.5	43.7	36.3	33.3	33.2	35.1

¹ The long-term unemployed are those people who have been unemployed for 27 weeks or longer.

² NZSCO: The New Zealand Standard Classification of Occupations of 1968 was revised in 1990, causing a discontinuity in the series in 1990/91.

³ MP&T = Managers, Professionals and Technicians.

⁴ S&SW = Services and Sales Workers.

⁵ A&FW = Agriculture and Fishery Workers.

⁶ Trades = Trade Workers.

⁷ P-MO&A = Plant and Machine Operators and Assemblers.

⁸ Elem. Occ. = Elementary Occupations.

Source: Statistics New Zealand, Labour Market Statistics 1999.

Table 3: Percentage Decomposition of Those Who Constitute the Long-term Unemployed¹ in New Zealand, 1987-1999

Variable	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Age Group													
15-19	30.3	25.8	18.3	17.2	16.3	15.4	13.7	13.0	12.1	14.2	12.6	11.9	11.9
20-24	13.1	20.0	22.2	19.3	17.8	20.8	19.1	18.2	17.0	13.4	15.2	16.2	14.7
25-34	23.8	22.6	26.1	26.2	27.3	26.9	27.0	25.7	24.4	21.6	22.8	22.6	20.4
35-44	14.8	14.8	16.9	16.1	19.4	18.1	19.6	20.2	21.4	24.9	21.3	23.8	24.9
45-54	12.3	11.0	9.9	15.0	13.2	13.4	14.0	16.3	17.6	17.7	19.2	17.8	18.2
+55	...	7.1	6.7	6.3	6.2	5.4	6.7	6.6	7.4	8.2	8.7	8.1	9.8
Gender													
Male	57.4	63.9	66.2	66.5	66.0	66.4	64.6	64.0	65.2	61.4	61.7	61.5	62.5
Female	42.6	36.1	33.8	33.5	34.0	33.6	35.4	36.0	34.8	38.6	38.3	38.5	37.5
Ethnic Group													
European	62.3	63.2	64.4	65.4	58.4 (79.1)	57.7	61.3	59.8	56.9	57.7 (69.0)	56.4	51.8	54.6
Maori	29.5	25.8	22.9	19.9	25.6 (9.7)	24.9	22.0	23.2	25.0	22.6 (14.5)	24.1	25.7	25.6
Pacific Isld	8.1	10.4	11.1 (3.8)	12.8	12.8	12.5	13.4	11.7 (5.6)	11.8	11.2	9.2
Other	3.9	3.8	4.7 (7.4)	4.6	3.8	4.6	4.7	7.7 (10.9)	7.6	11.4	10.8
Education Quals.													
No quals.	64.8	63.9	59.5	59.4	52.9	50.3	47.9	48.2	48.5	48.8	48.6	48.9	44.0
Sch. quals.	15.6	17.4	19.7	19.3	19.8	20.0	20.4	20.2	19.5	18.9	17.8	18.1	20.7
Post-sch.	9.0	9.0	9.9	11.4	11.8	11.3	11.8	13.3	13.5	12.2	11.8	11.4	12.7
Other quals	9.8	11.0	11.3	9.0	15.1	17.5	19.2	18.1	17.9	19.9	21.5	21.6	22.7
Occupation Sought²													
MP&T ³						8.4	9.4	9.2	8.3	10.9	10.0	11.9	11.9
Clerks						8.7	10.0	9.8	9.5	9.7	8.9	7.6	8.2
S&SW ⁴						10.8	12.1	11.8	9.3	10.4	11.8	12.1	12.1
A&FW ⁵						5.0	5.4	6.7	5.4	5.5	5.0	5.2	4.7
Trades ⁶						10.5	10.2	8.9	8.3	5.7	5.5	5.2	5.7
P-MO&A ⁷						7.1	6.8	7.9	6.9	6.5	7.1	6.7	6.3
Elem. Occ ⁸						15.2	20.4	25.3	23.7	14.7	15.0	15.9	12.1
Not specifd						34.3	25.3	20.2	28.6	36.6	37.0	35.4	39.3

Note: Figures may not add to 100.0 because of rounding. Figures in parentheses for ethnic groups show percentage shares in total population in the respective census years.

¹ The long-term unemployed are those people who have been unemployed for 27 weeks or longer.

² NZSCO: The New Zealand Standard Classification of Occupations of 1968 was revised in 1990 causing a discontinuity in the series in 1990/91.

³ MP&T = Managers, Professionals and Technicians.

⁴ S&SW = Services and Sales Workers.

⁵ A&FW = Agriculture and Fishery Workers.

⁶ Trades = Trade Workers.

⁷ P-MO&A = Plant and Machine Operators and Assemblers.

⁸ Elem. Occ. = Elementary Occupations.

Source: Statistics New Zealand, Labour Market Statistics 1999.

Important modifications have been made in the empirical logit model in this extended work. For example, age is modelled as a polychotomous rather than a continuous variable; the categories for ethnicity and educational attainment have been expanded; annual dummy variables have been utilised so that the full dataset can be analysed in one step; the geographical locations have been re-arranged to coincide with political regional council areas; and distributed lags of the growth rate of the economy have been successfully incorporated. Constrained by the set of available variables on which information is gathered by the NZES and the characteristically low predictive power of the statistical models, the parameters obtained in this study form a highly credible objective basis for estimating the probability an identified jobseeker has of suffering LTU. The *a posteriori* success rate of about 55% of the model in identifying LTU sufferers bodes well for the time when currently unavailable but relevant information on jobseekers and local labour markets can be integrated in the model.

The paper contains five sections. A short review of pertinent earlier studies is provided in Section 2. The dataset used and analytical method utilised are described in Section 3. Section 4 presents and discusses the empirical results. The last section provides a summary and outlines the conclusions from the study.

2. A REVIEW OF EMPIRICAL STUDIES ON UNEMPLOYMENT DURATION

2.1 General Introduction

In the literature, most empirical studies modelling the determinants of the length of out-of-work duration time for unemployed persons are based on survival analysis. The data for survival analysis arise in investigations where the subjects of study are followed until they reach a specified endpoint which is often, but by no means always, death (Everitt and Der 1996 p. 93). In unemployment duration models, survival time is the spell of unemployment, i.e., the time it takes for a job search to end in a successful [re-]employment. Because survival times are restricted to being positive, and because the data often contain censored observations (i.e., observations for which at the end of the study the event has not occurred) the familiar multiple regression techniques may not be justifiable. Central to the analysis of survival times are two functions that can be used to describe their distribution, namely the *survival function* and the *hazard function*.

The survival function $S(t)$ is defined as the probability that an individual survives longer than t , i.e., $S(t) = \Pr(T > t)$ where T denotes survival time. The graph of $S(t)$ versus t is known as the survival curve and it is generally negatively sloped. From such fitted curves, researchers can estimate the probability that an individual with a given set of characteristics will survive past a given length of time. On the other hand, the hazard function $h(t)$ is defined as the probability that an individual experiences the event of interest in a small time interval s , given that the individual has survived up to the beginning of this interval. The hazard function is also known as the instantaneous failure rate; it is a measure of how likely an individual is to experience an event at a particular point in time. Empirically, it is assumed that the survival and hazard functions are adequately captured by the Weibull distribution.

2.2 International Studies

Predicated on job search theory, most empirical research has been on the estimation of the hazard function in which the probability is made a function of different exogenous variables that usually include personal characteristics of those studied, conditions in the local labour market, the length of spell of unemployment and economic incentives such as the magnitude of the unemployment benefits (i.e., the replacement ratio). The parameters are usually estimated via the maximum likelihood approach and, even though the particular functional form may vary, the most frequently used specification is the proportional Weibull hazard function:

$$\lambda(t) = \alpha t^{\alpha-1} e^{x\beta} \cdot v$$

where t represents the length of the unemployment spell; α is the parameter used to capture the effect of possible duration dependence; v is a stochastic error component that is there to pick up heterogeneity from omitted variables, x and β are vectors of explanatory/independent variables and their coefficients, respectively (Holmlund *et al.*, 1989, p. 175). Values of $0 < \alpha < 1$ would suggest negative duration dependence; $\alpha = 1$, no duration dependence; and $\alpha > 1$, positive duration dependence. Variance of the error term $\sigma^2 \neq 0$ would suggest a heterogeneous sample, whilst $\sigma^2 = 0$ would suggest a fairly homogeneous sample.

Kaitz (1970) is one of the seminal papers on unemployment duration in the US. Lancaster (1979) was the first to suggest and empirically use the Weibull hazard function in a UK study. Other early studies include Nickell (1979a), Kiefer and Neuman (1979, 1981), Narenranathan *et al.* (1985) and Albrecht *et al.* (1986). Some of the important country studies include Clarke and Summers (1979) and Akerlof and Main (1980) for the US; Björklund (1981) for Sweden; Main (1981) and Nickell (1979b) for the UK; Hassan and De Broucker (1982) for Canada; and Trivedi and Baker (1983) for Australia. Good reviews of those studies have been done by Kiefer (1988), Holmlund *et al.* (1989) and Layard *et al.* (1991, Ch. 5).

Almost all the studies report negative duration dependence (i.e., exit probability or the probability of re-employment diminishes with length of unemployment spell) but the evidence on the effect of benefits on unemployment duration is mixed. In their survey of empirical studies, Holmlund *et al.* (1989) surmise that estimations of the benefit elasticity of unemployment duration based on time series data are not robust; those based on cross sectional data report values ranging from 0.3 to 0.6. Persons in age group 15-24 years are reported to have higher rates but shorter spells of unemployment than other groups in the population. The European studies show that personal characteristics (e.g., age, health, gender, marital status, and previous employment record) are more important than the labour market situation and the replacement ratio for expected spell lengths (Holmlund *et al.* 1989, pp. 209, 210 216).

The standard specification of the Weibull hazard function is not without flaws. For one thing, theoretical and empirical problems with the sensitivity of *ad hoc* parameterisations have led some writers to ponder what features of duration models can be identified non-parametrically. The seminal paper on “non-parametric techniques” is Cox (1972) that shows how to recover β without specifying the form of the baseline hazard, a process that does not allow the estimation of the duration dependence. To address that weakness and other data problems, Heckman and Singer

(1984) and Han and Hausman (1990), among others, have proposed various alternative extensions of Cox's model. A drawback of Han and Hausman's approach is the assumption of a prespecified distribution for the error term about which economic theory offers little guidance. Anton *et al.* (2001) have further extended Han and Hausman's approach by adopting an alternative semiparametric estimator of a proportional hazard model that has been proposed by Lewbel (2001), albeit in a different context. Using duration data constructed from a panel sample of unemployed men in Spain and focusing on the issues of heterogeneity and censoring, Anton *et al.* (2001) report that (i) the effect of censoring in a parametric model is to change the signs and magnitudes of the coefficients; (ii) heterogeneity is detected by using a censored ordered logit model with the modification proposed by Han and Hausman (1990) but the signs and magnitudes of the coefficients are not significantly affected; and (iii) if allowance is made for a wider class of heterogeneity, as done in their semiparametric model, then changes in both signs and sizes of the coefficients are observed.

2.3 New Zealand Studies

Among the empirical studies on New Zealand unemployment duration, Gardiner (1995) calculates the exit probability function (EPF) using the NZES data covering the period July 1989 to June 1993 and analyses the EPF to see if its shape and/or level changed over time and whether individuals with different characteristics have different exit probabilities. The EPF for 1993 is presented as a representative EPF and examined in detail to ascertain the different exit rates for the different demographic groups. The data contained censored observations and the approach taken in the exit probability analysis is tantamount to estimating a hazard function to generate the probability that an individual who has reached a specified duration will leave the unemployment register within the following seven days. It is reported that the EPF declines as duration increases, implying that the longer a person remains unemployed the less chance he or she has of leaving unemployment in the following week. Or it could also be that certain individual characteristics underlie a lower exit probability. The paper goes on to outline the disaggregated EPF according to gender, age group, ethnicity, educational qualifications and regional location (the variables on which NZES collects information). The conclusion, however, is that the model has low predictive power because of omitted relevant explanatory variables such as income (both from a previous job and from sources whilst unemployed), the number of dependents and local demand conditions.

Watson *et al.* (1997) take the "continuance curve" (or the survival curve) approach to investigate the relationship between jobseeker characteristics and the duration of unemployment spells in New Zealand. The dataset used comprised a little over 300,000 cases covering the period January 1995 to February 1997 and contained censored observations for which survival analysis is well-suited. The authors of that paper report that the continuance curve approach does not allow accurate identification of NZES jobseekers who are likely to become long-term unemployed on the basis of the few key demographic characteristics that are measured when the jobseekers enroll with the NZES. In predicting cases of LTU, the model always included a high proportion of short-term spells. The paper recommends that a broader range of jobseeker and local labour market information may improve the ability to model the duration of unemployment spells and to more accurately identify jobseekers who are most at risk of LTU.

Moving away from survival analysis, Obben *et al.* (2002) fitted a binary choice model to a random sample of 100,00 observations of completed unemployment spells covering a much longer period (1988-97) thereby covering different phases of the business cycle. The cases for individual years were analysed as well as data for the whole decade. The objective was to determine the probability that an unemployed individual with a given set of attributes will be long-term unemployed. The regression results made sense but the predictive power of the model was low.

3. ANALYTICAL METHOD AND DATA

The research objective is to estimate the extent to which the variables on which NZES collects information about jobseekers – age, gender, ethnicity, educational qualifications and regional location – are predictive of whether or not an unemployed individual would be long-term unemployed. The model of choice is the logit model⁵ in which the dependent variable is a dummy variable taking the value 1 if the spell of unemployment is 27 weeks or longer (i.e., a case of New Zealand-defined LTU) or 52 weeks or longer (i.e., a case of OECD-defined LTU) and 0 otherwise. Using the maximum likelihood estimation procedure, the discrete and dichotomous dependent variable Y_i is regressed on a set of discrete and/or continuous explanatory variables yielding a model that can be written as

$$(1) \quad Y_i = \ln [P_i / (1 - P_i)] = \hat{\alpha}_1 + \hat{\alpha}_2 X_{i2} + \dots + \hat{\alpha}_k X_{ik} \quad i = 1, \dots, N$$

where \ln is natural logarithm, P_i is the probability that individual i is a long-term unemployed person, the $\hat{\alpha}$'s are the regression coefficients and the X 's are the explanatory variables. Since the logit model assumes that Y_i is a logistic random variable, the dependent variable is the *log of the odds* that an individual will be a long-term unemployed. A given slope coefficient shows how the log of the odds (that an individual will be a long-term unemployed) changes as the corresponding explanatory variable changes by one unit (in the case of a continuous or discrete variable), or as an attribute different from that of the base category is considered (in the case of a dummy variable). The statistical significance of the slope coefficients may be assessed from their respective standard errors, t-ratios or p-values. A test of the hypothesis that all of the regression coefficients in the model are zero can be done via the likelihood ratio test where the chi-square test statistic has $k-1$ degrees of freedom. In dichotomous dependent variable models, the conventionally computed coefficient of determination (the R-square) is of questionable value as a measure of goodness of fit (Gujarati, 1995, pp. 545-546, 561, 579). Consequently, several alternatives are suggested in the literature. A number of these are reported with the analytical results in Section 4.

Furthermore, when the regression coefficients are exponentiated, the derived values or the antilogs indicate the effect of each explanatory variable directly on *the odds* of being a long-term unemployed rather than on the log-odds. Subtracting 1 from the antilogs and multiplying the results by 100 would give the *percentage changes in the odds* corresponding to one-unit change in the explanatory variables.

⁵ For the econometric theory of binary choice models see, for example, Intriligator (1978, pp. 173-176); and for the rationale for the preference of the logit over the probit model see, for example, Pindyck and Rubinfeld (1991, p. 256), Griffiths *et al.* (1993, p. 751) and Gujarati (1995, Ch. 16).

The *probability* that individual i would be long-term unemployed can be estimated from the antilogs of Equation (1):

$$(2) \quad P_i / (1 - P_i) = \text{antilog } \hat{\beta}_i \\ P_i = (1 - P_i) \text{antilog } \hat{\beta}_i \\ = \text{antilog } \hat{\beta}_i / (1 + \text{antilog } \hat{\beta}_i)$$

where $\hat{\beta}_i$ is the estimated value of the response variable from the regression for individual i .

The data for this study comprised all of the 2,476,898 cases of completed unemployment spells in the NZES database covering the period from 1988 to 1997. For each spell (determined from the start date and end date) the client number, age, gender, ethnicity, educational attainment and office of registration were noted. In preparing the data for analysis, the observations were sorted according to start date in order to note the number of cases in each year. Since there was no interest in the identities of the registrants, the client numbers were ignored. Cases containing no entry or meaningless entry or out-of-range values for any variable were deleted. All of the cases for the year 1997 were deleted because they were found to be a sub-selection of spells that ended in the same year; these were clearly incompatible with the others. After such corrections, the usable dataset came to 2,229,125 cases, implying about 10% of the original data were lost due to mistakes in them. Apart from the year 1988 which had 63,466 observations the other years had observations ranging from 252,152 to 285,272. The following statistics in weeks were noted about the unemployment spells: range: 0.14 to 477.14; mean: 31.03; median: 17.43; mode: 9.00; and standard deviation: 40.79. The distribution of the unemployment spells is certainly skewed to the right.

For estimation purposes the values of the various variables were divided into different categories so that dummy variables could be used. Unemployment spells 27 weeks or longer were given the dummy variable LTU27 to reflect the New Zealand definition of LTU; those 52 weeks or longer were given the dummy variable LTU52 to reflect the OECD definition of LTU. About 36% of the total cases were LTU27s, and about 16% were LTU52s.

The age variable was divided into nine classifications: TEENAGER (15-19 age group); EARLY20S (20-24 age group); LATE20S (25-29 age group); EARLY30S (30-34 age group); LATE30S (35-39 age group); EARLY40S (40-44 age group); LATE40S (45-49 age group); EARLY50S (50-54 age group); and BTWN5565 (55-65 age group). The base category for age was TEENAGER.

For obvious reasons, there were only two classifications for gender: MALE and FEMALE; the latter was used as the base category. The ratio of males to females in the dataset was 3:2.

The study employed six variables to match the six classifications of ethnicity reported in the NZES database: EUROPEAN (European or white); MAORI (Maori); MAO_EURO (mixed Maori-European); PACIFIC (Pacific Islander); MXDMAORI (mixed Maori-Pacific Islander); and OTHETHNC (all other ethnic groups). EUROPEAN was used as the base category for ethnicity.

There were eight educational attainment categories that the jobseekers are classified into; the corresponding variables are: LT3YRSCH (no formal school or less than three years of secondary

schooling); LT3SCERT (less than three School Certificate passes or equivalent); GE3SCERT (three or more School Certificate passes or equivalent); SXTIFORM (Sixth Form Certificate, University Entrance or equivalent); BURS_HSC (Bursary, High School Certificate); OSCHQUAL (other school or trade qualifications); POSTSECQ (post-secondary qualifications); and DEGPROFQ (university degree or professional qualifications). DEGPROFQ was utilized as the base category of educational attainment.

To reduce the number of points of geographical location, the study identified the jobseeker as being located in the regional council of the unemployment office or the city the jobseeker registered at. The variables used to represent the thirteen regional councils are as follows: NRTHLAND (Northland); AUCKLAND (Auckland); WAIKATO (Waikato); BOPLNTY (Bay of Plenty); HAWKSBBY (Hawke's Bay); TARANAKI (Taranaki); MANAWANG (Manawatu-Wanganui); WELLNTON (Wellington); NELSON (Nelson); WESCOAST (West Coast); CANTBBRY (Canterbury); OTAGO (Otago); SOUTHLND (Southland). Wellington, the national capital, was used as the base category.

4. EMPIRICAL RESULTS

Two different sets of the logit model were estimated. In the first set, Model 1, the NZES-based variables plus dummy variables for the various years were regressed on LTU27 and on LTU52. The annual dummy variables are meant to capture all the year-specific factors not reflected in the other variables. This model is principally to furnish *a posteriori* probabilities for the registered jobseekers and thereby provide a measure of the predictive power of the logit model. However, for *ex ante* probabilities, which are more relevant for ALMPs, the results of Model 1 cannot be used to estimate probabilities for either the current year or future years. To get around that problem, the model was re-estimated with distributed lags of the growth rate of the economy replacing the dummy variables for the individual years. Up to four lags were deemed adequate. The thinking is that when appropriate sets of growth figures (actual or predicted) are substituted along with the attributes of an identified jobseeker, the probability of LTU for that individual can be calculated for out-of-sample years. The regression results and the figures for *percentage change in odds* obtained for Model 1 and for Model 2 are reported in Table 4 and Table 5, respectively. The econometric software SHAZAM (White, 1997) was utilised for the computations. Since the main interest is in the implications of the New Zealand definition of LTU, the discussion will concentrate on the LTU27 results; the implications of the OECD definition of LTU may be inferred from an examination of the LTU52 figures and the discussion of the LTU27 results.

The LTU27 results in Table 4 show that only four (OTHRSCHQ, WAIKATO, OTAGO and DV94) out of the 43 regressors are not statistically significant. Even though the sampled R-squares are expectedly low, the chi-square score for the likelihood ratio test clearly leads to the rejection of the joint test that all the regression coefficients are equal to zero.

Table 4
Regression Results of Model 1

Variable	LTU27			LTU52		
	Coefficient	T-ratio	% Change in Odds	Coefficient	T-ratio	% Change in Odds
MALE	0.37645	127.32	45.71	0.43425	107.13	54.38
EARLY20S	0.21385	51.58	23.84	0.15259	27.05	16.48
LATE20S	0.14525	30.97	15.63	0.16972	27.01	18.49
EARLY30S	0.14811	28.46	15.96	0.22495	32.68	25.22
LATE30S	0.15264	26.41	16.49	0.25736	33.83	29.35
EARLY40S	0.16478	25.87	17.91	0.32356	39.17	38.2
LATE40S	0.23701	33.23	26.74	0.43297	47.72	54.18
EARLY50S	0.32966	40.06	39.05	0.55622	55.01	74.41
BTWN5565	0.22267	21.25	24.94	0.06746	4.57	6.97
MAORI	0.26754	67.87	30.67	0.34095	67.94	40.62
MAO_EURO	0.13096	21.83	13.99	0.16219	20.41	17.61
PACIFIC	0.21919	36.27	24.51	0.29207	38.24	33.92
MXDMAORI	0.14785	10.45	15.93	0.20376	10.98	22.6
OTHETHNC	0.10361	14.91	10.91	0.12418	12.99	13.22
LT3YRSCH	0.32183	49.93	37.96	0.50594	54.06	65.85
LT3SCERT	0.16582	23.15	18.03	0.29341	28.47	34.1
GE3SCERT	0.07469	10.29	7.75	0.16425	15.61	17.85
SXTHFORM	0.05372	7.27	5.52	0.10006	9.23	10.52
BURS_HSC	0.12773	13.64	13.62	0.16361	11.89	17.77
OTHRSCHQ	0.00639	0.49	0.64	0.05336	2.86	5.48
POSTSECQ	-0.03655	-4.43	-3.58	0.03177	2.67	3.23
NRTHLAND	0.10653	12.99	11.24	0.17303	16.67	18.89
AUCKLAND	-0.01561	-2.81	-1.54	0.03259	4.457	3.31
WAIKATO	-0.00393	-0.58	-0.39	0.00864	0.98	0.87
BOPLENTY	0.01501	2.16	1.51	0.01902	2.11	1.92
HAWKSBAY	-0.07113	-10.43	-6.86	-0.24599	-26.72	-21.81
TARANAKI	-0.06848	-7.51	-6.62	-0.09321	-7.63	-8.9
MANAWANG	-0.04802	-6.23	-4.68	0.01249	1.23	1.25
NELSON	-0.07587	-9.49	-7.31	-0.19707	-17.91	-17.88
WESCOAST	-0.07637	-5.21	-7.35	-0.15785	-7.86	-14.6
CANTBURY	-0.10799	-16.93	-10.23	-0.17273	-19.93	-15.86
OTAGO	-0.00801	-0.98	-0.79	-0.06451	-5.85	-6.25
SOUTHLND	-0.19683	-20.85	-17.86	-0.35755	-27.26	-30.06
DV89	-0.03797	-4.01	-3.72	-0.00114	-0.09	-0.11
DV90	0.18174	19.42	19.93	0.07888	6.72	8.21
DV91	0.44961	48.29	56.77	0.35641	30.79	42.82
DV92	0.35578	38.12	42.73	0.09454	8.05	9.91
DV93	0.18816	20.11	20.71	-0.14276	-12.01	-13.3
DV94	-0.01287	-1.36	-1.28	-0.38393	-31.49	-31.88
DV95	-0.08876	-9.33	-8.49	-0.47085	-38.19	-37.55
DV96	-0.21609	-22.55	-19.43	-0.97374	-74.29	-62.23
CONSTANT	-1.27976	-109.29	-72.19	-2.40323	-153.56	-90.95
<i>Summary statistics</i>						
Likelihood Ratio Test	65991 with 41 d.f.			81272 with 41 d.f.		
Maddala R ²	0.029			0.036		
Cragg-Uhler R ²	0.040			0.062		
McFadden R ²	0.023			0.042		
Prediction success rate	55.1%			74.5%		

Table 5
Regression Results of Model 2

Variable	Dependent variable is LTU27			Dependent Variable is LTU52		
	Coefficient	T-ratio	% Change in Odds	Coefficient	T-ratio	% Change in Odds
MALE	0.37663	127.46	45.73	0.43581	107.57	54.62
EARLY20S	0.21572	52.07	24.07	0.15384	27.29	16.63
LATE20S	0.14805	31.59	15.95	0.17186	27.37	18.75
EARLY30S	0.15123	29.09	16.32	0.22597	32.85	25.35
LATE30S	0.15595	26.99	16.87	0.25815	33.95	29.45
EARLY40S	0.16844	26.46	18.34	0.32471	39.33	38.36
LATE40S	0.24091	33.79	27.24	0.43371	47.82	54.29
EARLY50S	0.33387	40.61	39.63	0.55704	55.12	74.55
BTWN5565	0.22831	21.79	25.64	0.06347	4.3	6.55
MAORI	0.26626	67.59	30.51	0.33832	67.45	40.26
MAO_EURO	0.13509	22.53	14.46	0.16454	20.71	17.88
PACIFIC	0.21805	36.11	24.36	0.28986	37.97	33.62
MXDMAORI	0.14815	10.47	15.96	0.19953	10.76	22.08
OTHETHNC	0.10486	15.09	11.05	0.12145	12.71	12.91
LT3YRSCH	0.32246	50.05	38.05	0.50915	54.43	66.38
LT3SCERT	0.16789	23.45	18.28	0.29706	28.84	34.59
GE3SCERT	0.07821	10.78	8.13	0.16971	16.13	18.49
SXTHFORM	0.05734	7.76	5.91	0.10417	9.61	10.98
BURS_HSC	0.13132	14.03	14.03	0.16641	12.11	18.11
OTHRSCHQ	0.00867	0.68	0.87	0.05711	3.06	5.87
POSTSECQ	-0.03593	-4.36	-3.53	0.03332	2.81	3.39
NRTHLAND	0.10719	13.08	11.31	0.17326	16.7	18.91
AUCKLAND	-0.01418	-2.55	-1.41	0.03439	4.71	3.5
WAIKATO	-0.00493	-0.74	-0.49	0.00899	1.02	0.9
BOPLENTY	0.01353	1.95	1.36	0.01986	2.2	2.01
HAWKSBAV	-0.07288	-10.69	-7.03	-0.24662	-26.8	-21.85
TARANAKI	-0.07038	-7.72	-6.79	-0.09436	-7.73	-9
MANAWANG	-0.04807	-6.24	-4.69	0.01135	1.12	1.14
NELSON	-0.07678	-9.61	-7.39	-0.19583	-17.8	-17.78
WESCOAST	-0.07478	-5.11	-7.21	-0.15719	-7.83	-14.54
CANTBURY	-0.10632	-16.68	-10.08	-0.17174	-19.83	-15.78
OTAGO	-0.00916	-1.13	-0.91	-0.06404	-5.81	-6.2
SOUTHLND	-0.19522	-20.1	-17.73	-0.35379	-26.99	-29.79
GRTHRATE	-0.07957	-90.84	-7.65	-0.11634	-99.39	-10.98
GRATEL1	0.01738	15.96	1.75	0.04529	31.24	4.63
GRATEL2	-0.06977	-65.29	-6.74	-0.13754	-91.37	-12.85
GRATEL3	-0.02269	-17.05	-2.24	0.00741	4.44	0.74
GRATEL4	-0.03878	-41.16	-3.81	0.03357	27.62	3.41
CONSTANT	-0.94152	-110.84	-60.99	-2.32328	-193.94	-90.2
<i>Summary statistics</i>						
Likelihood Ratio Test	63776 with 38 d.f.			79945 with 38 d.f.		
Maddala R ²	0.028			0.035		
Cragg-Uhler R ²	0.039			0.061		
McFadden R ²	0.022			0.041		
Prediction success rate	55.0%			74.5%		

The figures for percentage change in odds suggest that males are 46% more likely than females to suffer LTU. This result is in consonance with the finding by the Employment Task Force which reported that women are more likely to “leave the labour force” on becoming unemployed (1994, p. 14). All the age groups are more likely than teenagers to suffer LTU. Ranking them, the logit results suggest that those in the early 50s are the worst off, followed by those in their late 40s, 55 and older, early 20s, early 40s, late 30s, early 30s and late 20s. This is consistent with much of what is reported in the literature.

Compared to those identified as Europeans/whites, all the other races are more likely to suffer LTU. The Maori are 31% more likely than Europeans to suffer LTU, followed by Pacific Islanders (25%), mixed Maori-Pacific Islanders (16%), mixed Maori-European (14%) and other ethnic groups (11%). These figures tell the same qualitative story as the figures in Tables 2 and 3.

With respect to educational attainment, the only group that does better than university graduates are those with post-secondary qualifications – they are 4% less likely than graduates to suffer LTU. Those with “other school qualifications” are nearly at par with graduates. As is to be expected, those with no schooling or less than three years of schooling are the worst off, being 38% more likely than graduates to suffer LTU. They are followed by those with less than three school certificates (18%), those with bursary or high school certificate (14%) and those with sixth form qualifications (6%).

Being the most economically depressed part of the country, Northland unsurprisingly turns out the highest positive percentage change in odds among the regional locations. The only other place to turn out a positive figure is Bay of Plenty. Whereas registrants in Northland are 11% more likely to suffer LTU than those in Wellington, the figure for Bay of Plenty is only 2%. The rest of the regions range from slightly better to significantly better than Wellington. The region least likely to experience LTU is Southland, followed by Canterbury, West Coast, Nelson, Hawke’s Bay, Taranaki, Manawatu-Wanganui and Auckland.

Using 1988 as the base year, the logit results indicate that everyone’s odds of experiencing LTU decreased by 4% in 1989 but increased by 20%, 57%, 43% and 21%, respectively, in the next four years. Subsequently, the odds decreased by 1%, 8% and 19% in 1994, 1995 and 1996, respectively. It may be noted that New Zealand experienced acute recession in both 1988 and 1989.

When the annual dummy variables are replaced by the distributed lags of the growth rate of the economy, the LTU27-results of Model 2 in Table 5 show that the coefficients for the demographic and regional variables are comparable to those in Model 1, suggesting that the parameters from the logit model are quite robust. Turning attention to the growth variables, it can be seen that all of them have highly significant coefficients. The configuration of their signs suggests that an increase in the growth rate of the economy may decrease the log of the odds of LTU in the current year, raise it in the second year and lower it continually thereafter. The second-year effect is puzzling and would require investigation.

Finally, the parameters of Model 2 will be used to illustrate the estimation of out-of-sample probabilities of LTU for a few representative jobseekers. The March 2002 publication of Key Statistics [for New Zealand] shows the March-ending annual growth rates to be 3.06%, 1.85%,

0.41%, 4.68% and 2.67% for the years 1997 through 2001. These growth rates have been used in conjunction with the LTU27-parameters of Model 2 to predict the values of the dependent variable for eight random individuals and to calculate their respective probabilities of suffering LTU in the year 2001. Individual 'A' is a male, mixed Maori-European in his late 20s having other school qualifications and living in Auckland. When the location of individual 'A' is changed to Northland, he is identified as individual 'B'. Individuals 'C' and 'D' are the European/white counterparts of 'A' and 'B', respectively. The female equivalents of 'A', 'B', 'C' and 'D' are named 'E', 'F', 'G' and 'H'. The values of the response variable for these individuals and the calculated probabilities of LTU are summarized in Table 6.

Table 6
Estimating Out-of-Sample Probability of LTU for Representative Jobseekers

Individual	Estimated Value of the Response Variable (= Log of the Odds of Experiencing LTU)	Probability of LTU in 2001
A	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.37663 + 0.14805 + 0.13509 + 0.00867 - 0.01418 = -0.60762$	35.26%
B	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.37663 + 0.14805 + 0.13509 + 0.00867 + 0.10719 = -0.48625$	38.08%
C	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.37663 + 0.14805 + 0.00867 - 0.01418 = -0.74271$	32.24%
D	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.37663 + 0.14805 + 0.00867 + 0.10719 = -0.62134$	34.95%
E	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.14805 + 0.13509 + 0.00867 - 0.01418 = -0.98425$	27.20%
F	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.14805 + 0.13509 + 0.00867 + 0.10719 = -0.86288$	29.67%
G	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.14805 + 0.00867 - 0.01418 = -1.11934$	24.61%
H	$-0.94152 - (0.07957*2.67) + (0.01738*4.68) - (0.06977*0.41) - (0.02269*1.85) - (0.03878*3.06) + 0.14805 + 0.00867 + 0.10719 = -0.99797$	26.93%

Note: definitions of individuals A to H are given in the text.

As it will seen in the last column of Table 6, changing the location of individual 'A' from Auckland to Northland increases his probability of experiencing LTU from 35% to 38%. If the person's ethnicity was not mixed Maori-European but European/white, his probability of experiencing LTU

would be three percentage points lower at both places. Changing the gender to female reduces the respective probabilities by eight percentage points but the relative ethnic and locational differences are maintained. It would seem from this illustration that plausible measures of the probability of suffering LTU can be obtained for all identifiable jobseekers. Given that there are $2*8*6*13*9 = 11,232$ possible combinations of the attributes of registrants from the classifications of gender, educational attainment, ethnicity, regional office and age group, the illustration above should suffice. The probabilities calculated this way can constitute an objective basis which can be augmented with relevant supplementary information to decide on the suitability of a particular ALMP for various jobseekers.

5. SUMMARY AND CONCLUSION

To advance a statistical profiling model for New Zealand, this paper has extended a previous work based on a random sample of 100,000 to cover more than 2.2 million cases of completed unemployment spells over the 1988-97 period. Different variants of a logit model were estimated to find the relationship between a set of recorded attributes describing a registered jobseeker and the probability that the individual would be long-term unemployed. The signs and magnitudes of the regression coefficients and other measures calculated from them lead to the endorsement of what is qualitatively known about the demographic and locational characteristics of the unemployed in the country. The fit of the model is satisfactory but the predictive power is characteristically low. However, the incorporation of the distributed lags of the growth rate of the economy is an innovation that can allow the calculation of out-of-sample probability of suffering LTU for any identifiable jobseeker. Consequently, this study advocates the adoption of the parameters obtained herein as the first step in the construction of a formal profiling model for New Zealand.

REFERENCES

- Akerlof, G. A. and Main, B. G. (1980) Unemployment Spells and Unemployment Experience, *American Economic Review*, Vol 70, No. 5, pp. 885-893.
- Albrecht, J. W., Holmlund, B., and Lang, H. (1986) Job Search and the Transition to Employment: Theory, FIEF Working Paper No. 26, Stockholm.
- Anton, A. A., Sainz, A. F. and Rodriguez-Poo, J. (2001) Semiparametric Estimation of a Duration Model, *Oxford Bulletin of Economics and Statistics*, Vol 63, No. 5, pp. 517-533.
- Barron, J. M. and Mellow, W. (1981) Changes in Labor Force Status Among the Unemployed, *Journal of Human Resources*, Vol 16.
- Björklund, A. (1981) Studies in the Dynamics of Unemployment, EFI (Economic Research Institute), Stockholm.
- Clark, K. B. and Summers, L. H. (1979) Labor Market Dynamics and Unemployment: A Reconsideration, *Brookings Papers on Economic Activity*, Vol 1, pp.13-60.
- Cox, D. R. (1972) Regression Models and Life-Tables, *Journal of Royal Statistical Society, Ser. B.*, Vol 26, pp. 186-220.
- Cross, R. (1995) The Natural Rate of Unemployment: Reflections on 25 Years of the Hypothesis (ed), Cambridge University Press, Cambridge.
- Employment Task Force (1994) Employment: The Issues, Prime Ministerial Task Force on Employment, Office of the Prime Minister, Wellington.
- Everitt, B. S. and Der, G. (1996) A Handbook of Statistical Analyses Using SAS, Chapman & Hall, London.
- Friedman, M (1968) The Role of Monetary Policy, *American Economic Review*, Vol 58, No. 1, pp. 1-17.
- Galbraith, J. K. (1997) Time to Ditch the NAIRU, *Journal of Economic Perspectives*, Vol 11, No. 1, pp. 93-108.
- Gardiner, P. (1995) "An Analysis of Exit Rates and Duration Dependence in Registered Unemployment" in Morrison, P. S. (ed.) (1995) *Labour, Employment and Work in New Zealand 1994*, Proceedings of the Sixth Conference 24-25 November 1994, Victoria University of Wellington.
- Gordon, R. (1997) The Time-Varying NAIRU and its Implications for Economic Policy, *Journal of Economic Perspectives*, Vol 11, No. 1, pp. 11-32.
- Griffiths, W. E., Hill, R. C. and Judge, G. G. (1993) Learning and Practicing Econometrics, John Wiley & Sons, New York.
- Gujarati, D. N. (1995) Basic Econometrics, 3rd edition, McGraw-Hill, Inc., New York.
- Hall, R. E. (1979) A Theory of the Natural Rate of Unemployment and the Duration of Unemployment, *Journal of Monetary Economics*, Vol 5, pp. 153-169.
- Han, A. and Hausman, J. (1990) Flexible Parametric Estimation of Duration and Competing Risk Models, *Journal of Applied Econometrics*, Vol 5, pp. 1-28.
- Hasan, A. and De Broucker, P. (1982) Duration and Concentration of Unemployment, *Canadian Journal of Economics*, Vol 15 No. 4, pp. 735-756.
- Heckman, J. and Singer, B. (1984) A Method for Minimizing the Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, Vol 52, pp. 271-320.
- Heikensten, L. (1984) Studies in Structural Change and Labor Market Adjustment, EFI (Economic Research Institute) Stockholm.

- Holmlund, B., Löfgren, K. G. and Engström, L. (1989) Trade Unions, Employment and Unemployment Duration, Clarendon Press, Oxford.
- Intriligator, M. D. (1978) Econometric Models, Techniques and Applications, Prentice-Hall, Inc., Englewood Cliffs, N.J., U.S.A.
- Kaitz, H. B. (1970) Analyzing the Length of Spells of Unemployment, *Monthly Labor Review*, Vol 93, pp. 11-20.
- Katz, L. F. (1994) Active Labour Market Policies to Expand Employment and Opportunity, Federal Reserve Bank of Kansas City, Symposium on Reducing Unemployment: Current Issues and Policy Options, pp. 239-290.
- Kiefer, N. M. (1988) Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, Vol 26.
- Kiefer, N. M. and Neuman, G. R. (1979) An Empirical Job Search Model With a Test of the Constant Reservation Wage Hypothesis, *Journal of Political Economy*, Vol 87, pp. 89-107.
- Kiefer, N. M. and Neuman, G. R. (1981) Individual Effects in a Nonlinear Model: Explicit Treatment of Heterogeneity in the Empirical Job-Search Model, *Econometrica*, Vol 49 No. 4, pp. 965-979.
- Key Statistics (various issues), Statistics New Zealand, Wellington, New Zealand.
- Lancaster, T. (1979) Econometric Methods for the Duration of Unemployment, *Econometrica*, 47, pp. 939-956.
- Layard, R., Nickell, S. and Jackman R. (1991) Unemployment: Macroeconomic Performance and the Labour Market, Oxford University Press, Oxford.
- Lewbel, A. (2000) Semiparametric Qualitative Response Model Estimation with Unknown Heteroskedasticity or Instrumental Variables, *Journal of Econometrics*, Vol 97, pp. 145-177.
- Main, B. G. (1981) The Length of Employment and Unemployment in Great Britain, *Scottish Journal of Political Economy*, Vol 28, pp. 14-164.
- Martin, J. P. (1994) The Extent of High Unemployment in OECD Countries, Federal Reserve Bank of Kansas City, Symposium on Reducing Unemployment: Current Issues and Policy Options, pp. 5-40.
- Narendranathan, W., Nickell, S. and Stern, J. (1985) Unemployment Benefits Revisited, *Economic Journal*, Vol 95, pp. 307-329.
- Nickell, S. (1979a) Estimating the Probability of Leaving Unemployment, *Econometrica*, Vol 47 No. 5, pp.1249-1266.
- Nickell, S. (1979b) The Effect of Unemployment and Related Benefits on the Duration of Unemployment, *Economic Journal*, Vol 89, pp. 34-49.
- Nickell, S. (1990) Unemployment: A Survey, *Economic Journal*, Vol 100, No. 401, pp. 391-439.
- Obben, J., Engelbrecht, H-J. and Thompson, V. W. (2002) A Logit Model of the Incidence of Long-Term Unemployment, *Applied Economics Letters*, Vol 9, pp. 43-46.
- OECD Proceedings (1998) Early Identification of Jobseekers at Risk of Long-Term Unemployment: The Role of Profiling, OECD Secretariat, Paris.
- Phelps, E. S. (1967) Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time, *Economica*, Vol 34, No. 3, pp. 254-281.
- Pindyck, R. S. and Rubinfeld, D. L. (1991) Econometric Models and Economic Forecasts, 3rd edition, McGraw-Hill, Inc., New York.

- Richardson, P., Boone, L., Giorno, C., Meacci, M., Rae, D. and Turner, D. (2000) The Concept, Policy Use and the Measurement of Structural Unemployment: Estimating a Time-Varying NAIRU Across 21 OECD Countries, OECD Economics Department Working Papers No. 250.
- Salant, S. W. (1977) Search Theory and Duration Data: A Theory of Sorts, *Quarterly Journal of Economics*, Vol 91, pp. 38-57.
- Staiger, D., Stock, J. H. and Watson, M. W. (1997) The NAIRU, Unemployment and Monetary Policy, *Journal of Economic Perspectives*, Vol 11, No. 1, pp. 33-49.
- Stiglitz, J. (1997) Reflections on the Natural Rate Hypothesis, *Journal of Economic Perspectives*, Vol 11, No. 1, pp. 3-10.
- Trivedi, P. K. and Baker, G. M. (1983) Unemployment in Australia: Duration and Recurrent Spells, *Economic Record*, Vol 59, pp. 132-148.
- Watson, R., Maré, D. and Gardiner, P. (1997) Predicting the Duration of Unemployment Spells, *Labour Market Bulletin* (New Zealand Dept of Labour) Vol 2, pp. 51-65.
- White, K. (1997) SHAZAM: User's Reference Manual Version 8, McGraw-Hill Book Company, New York.