



TI 2000-030/3
Tinbergen Institute Discussion Paper

The Non-Parametric Identification of Lagged Duration Dependence

Paul Frijters

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>

The non-parametric identification of lagged duration dependence.

Paul Frijters*

March 16, 2000

Abstract

In this paper, I consider the identification of lagged duration dependence in multiple spells without using the assumption that there are additional regressors orthogonal to the individual effects. The non-parametric identification strategy is applied to the multiple non-employment spells of 2066 new entrants. The estimated lagged duration dependence function rises for the first 4 months and thereafter

*Corresponding Author. Department of Economics, Room 2A-42, Free University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. E-mail: pfrijters@econ.vu.nl. Ph: +31-20-4446155.

decreases, suggesting negative effects of non-employment for long periods of non-employment.

YEL-code: C41, C14, J64

1 Introduction

In this paper, the identification of lagged duration dependence in multiple spells is considered theoretically and empirically. The main interest is to examine the claims of some of the theories in the hysteresis literature (see Roed, 1995) that hazard rates into employment decrease if previous spells of non-employment were longer. The extension to the known results by Honoré (1993), Ridder (1991) and Heckman (1990), is that identification is shown without using the assumption that there are additional regressors orthogonal to the individual effects. The ‘cost’ of this result is that we need to assume that baseline functions are constant in the period under consideration.

After showing that the model is identified and after discussing some extensions, the model is estimated with a longitudinal administrative Dutch Income Panel data set. From the 65,000 individuals followed over the period 1989-1997, we focus on 2066 individuals who entered the labour market during the period under consideration and have multiple spells. The result-

ing lagged duration dependence function rises for the first 4 months and thereafter decreases, suggesting negative effects of non-employment for long periods of non-employment. These results are somewhat in line with those of Omori (1997) and Trivedi and Alexander (1989).

2 Lagged duration dependence in an MPH-model

Suppose the hazard rate in an MPH-model has the form:

$$\theta(t_{i,m}|t_{i,m-1}, \lambda_i) = \lambda_i z'(t_{i,m}) h(t_{i,m-1}) \quad (1)$$

whereby θ denotes the hazard rate, $t_{i,m}$ denotes the time that individual i has so far spent in the current state in the m^{th} spell of that state, and $h(t_{i,m-1})$ denotes the lagged duration dependence function which depends on time spent in the previous spell of the same state. The function $z(t_{i,m})$ denotes the baseline hazard. Denote the cumulative distribution of the individual effects λ_i as G .

This model is different from the ones discussed in the literature in one important respect: there are no individual regressors x which are independent of the individual effects. Such individual regressors are the main vehicle for proving identification in the one-spell models of Heckman and Singer (1984, RES), Ridder (1990), and Heckman (1991). Also in the identification results for multiple spells by Honoré (1993, RES) or Heckman and Borjas (1980), the models which allow for the hazard to depend on the length of the previous spell, assume a set of independent regressors x which help to identify the model. When individual observed characteristics are correlated with individual effects, the question arises which assumptions are required to identify the model in equation (1).

In the first period of new entrants into the labour market, there is no previous spell and the lagged duration for the first spell is hence 0. In terms of hazard rates, the hazard rates for the first two periods of new entrants then equals:

$$\begin{aligned}\theta_1(t_1|\lambda) &= \lambda h(0)z'(t_1) \\ \theta_2(t_2|\lambda, t_1) &= \lambda h(t_1)z'(t_2)\end{aligned}\tag{2}$$

If $h(\cdot)$ picks up loss-of-skill effects in the vein of Lungqvist and Sargent (1998), then new entrants who have not been non-employed before cannot have lost skills yet, and their previous duration of non-employment is zero.

By observing (t_1, t_2) for an infinite number of individuals, the joint and marginal distributions of t_1 and t_2 are known. Consider the following assumption:

assumption 1: $z(t)$ is a monotonically increasing continuous differentiable function with $z(0)=0$ and $z(\infty)=\infty$.

Under this assumption, the model is identified:

Theorem 1 *Under assumption 1, the function $h(\cdot)$, $z(\cdot)$ and the distribution G of the model in equation (2) are non-parametrically identified up to two scale-normalisations.*

Proof:

Denote the unconditional distribution of t_1 as $f(t_1)$. The following functions will be needed:

$$\begin{aligned}
f(t_1) &= \int \lambda h(0) z'(t_1) e^{-\lambda h(0) z(t_1)} dG = h(0) z'(t_1) \mathcal{L}'_G[h(0) z(t_1)] \\
dG(\lambda|t_1) &= \frac{\frac{\lambda h(0) z'(t_1) e^{-\lambda h(0) z(t_1)}}{\int \lambda h(0) z'(t_1) e^{-\lambda h(0) z(t_1)} dt_1}}{\int \frac{\lambda h(0) z'(t_1) e^{-\lambda h(0) z(t_1)}}{\int \lambda h(0) z'(t_1) e^{-\lambda h(0) z(t_1)} dt_1} dG} dG = \frac{\frac{\lambda h(0) z'(t_1) e^{-\lambda h(0) z(t_1)}}{1 - e^{-\lambda h(0) z(\infty)}}}{\int \frac{\lambda h(0) z'(t_1) e^{-\lambda h(0) z(t_1)}}{1 - e^{-\lambda h(0) z(\infty)}} dG} dG \\
&= \frac{\lambda e^{-\lambda h(0) z(t_1)}}{\mathcal{L}'_G[h(0) z(t_1)]} dG
\end{aligned}$$

and

$$\begin{aligned}
f(t_2|t_1) &= \int \lambda z'(t_2) h(t_1) e^{-\lambda z(t_2) h(t_1)} dG(\lambda|t_1) \\
&= \frac{z'(t_2) h(t_1)}{\mathcal{L}'_G[h(0) z(t_1)]} \mathcal{L}''_G[z(t_2) h(t_1) + h(0) z(t_1)] \\
S(t_2|t_1) &= \frac{\mathcal{L}'_G[z(t_2) h(t_1) + h(0) z(t_1)]}{\mathcal{L}'_G[h(0) z(t_1)]} \\
f(t_1, t_2) &= z'(t_2) h(t_1) h(0) z'(t_1) \mathcal{L}''_G[z(t_2) h(t_1) + h(0) z(t_1)]
\end{aligned}$$

where \mathcal{L} denotes the Laplace transform. The basic idea of the identification is that the distribution of t_2 conditional on t_1 identifies the baseline hazard and the function $h(\cdot)$, which in turn allows us to identify G . There holds:

$$\frac{\frac{f(t_1=t)}{\lim_{t_1 \downarrow 0} S(t_2=t|t_1)}}{f(t_1=0)} = \frac{z'(t)}{z'(0)}$$

and

$$\frac{\lim_{t_2 \downarrow 0} f(t_1=t, t_2)}{\lim_{t_1 \downarrow 0} f(t_1, t_2=t)} = \frac{h(t)}{h(0)}$$

Hence, we can identify $z'(t_1)$ up to the scalar $z'(0)$, and we can identify $h(t)$ up to the scalar $h(0)$. With the functions $z(t)$ and $h(t)$, we can find the distribution G from the uniqueness of the Laplace transform. We can hence find all the functions uniquely if we choose two out of the three values of $E\lambda$, $h(0)$ and $z'(0)$. Indeed, it can already be seen from the two hazard functions that the set $\{\lambda, h(\cdot), z(\cdot)\}$ generates observations indistinguishable from those generated by $\{a\lambda, bh(\cdot), \frac{1}{ab}z(\cdot)\}$ where a and b are positive constants. \square

We can apply this proof also to models incorporating cumulative effects. Consider for instance the following specification of the hazard rate:

$$\theta(t_{i,m}|t_{i,m-1}, \dots, t_1, \lambda_i) = \lambda_i z'(t_{i,m}) \prod_{j=1}^m h(t_{i,m-j})$$

For the pair of spells $\{t_{i,m}, t_{i,m-1}\}$, we can then define a new individual effect η_i , which is defined as $\eta_i \equiv \prod_{j=2}^m \lambda_i h(t_{i,m-j})$. By interchanging t_m with t_2 and t_{m-1} with t_1 , we then get the same model as defined in equation (2) with a different individual effect and with $h(0) = 1$. In fact, we can apply this procedure to all the pairs of observations $\{t_{i,j}, t_{i,j-1}\}$, $j=1, \dots, m$, of individual i and use all these pairs to identify $h(\cdot)$ and $z(\cdot)$ ¹. With this kind of model, we can trace the cumulative effect of all previous durations on current hazard rates.

Can we allow for different baseline hazards in each spell? Consider the following specification

¹There is then overidentifying information, because the individual effects for the set of pairs $\{t_j, t_{j-1}\}$, equals the individual effects for the set of pairs $\{t_{j-1}, t_{j-2}\}$ multiplied by $h(t_{j-2})$. Given that $h(\cdot)$ is also identified from $f(t_j, t_{j-1})$ and $f(t_{j-1}, t_{j-2})$ separately, this in principle allows for a specification check.

$$\theta_1(t_1|\lambda) = \lambda h(0)z'(t_1) \tag{3}$$

$$\theta_2(t_2|\lambda, t_1) = \lambda h(t_1)\tilde{z}'(t_2)$$

with $\tilde{z}'(t_2)$ and $z'(t_2)$ two different baseline functions. Without the lagged duration dependence function, Honoré (1993) showed that this model is identified. With the lagged duration dependence function, $t_1|\lambda$ is not independent of t_2 in the second spell however and the arguments of Honoré do not apply. Take the following assumptions:

assumption 2. all moments of G are finite.

assumption 3. $\lim_{t_1 \downarrow 0} z^k(t_1)$ and $\lim_{t_1 \downarrow 0} h^k(t_1)$ exist and are finite for all $k \in N^+$.

The following now applies:

Theorem 2 *The model above is identified up to scale normalizations under assumptions 2 and 3.*

proof: pick any positive $z'(0)$, $\tilde{z}'(0)$ and $h(0)$. Then, $E\lambda$ is known from

$\lim_{t_1 \downarrow 0} f(t_1) = h(0)z'(t_1)E\lambda$. We have

$$\begin{aligned} \frac{d^k f(t_1)}{d^k t_1} &= h(0)\{z^{k+1}(t_1)\mathcal{L}'_G[h(0)z(t_1)] + kz^k(t_1)h(0)\mathcal{L}''_G[h(0)z(t_1)] \\ &\quad + \dots + z'(t_1)^{k+1}\mathcal{L}_G^{k+1}[h(0)z(t_1)]\} \\ \frac{d^k f(t_1)}{d^k t_1}\Big|_{t_1=0} &= h(0)\{z^{k+1}(0)E\lambda + kz^k(0)h(0)^2E\lambda^2 + \dots + z'(t_1)^{k+1}h(0)^{k+1}E\lambda^{k+1}\} \end{aligned}$$

Hence, $\frac{d^k f(t_1)}{d^k t_1}$ depends on all the 1 to k+1 derivatives of $z(\cdot)$ and \mathcal{L}_G .

This implies that if all the derivatives of $z(\cdot)$ and \mathcal{L}_G up to and including the k'th derivatives are known and either $z^{k+1}(0)$ or $\mathcal{L}_G^{k+1}[0]$ is known, that the remaining $z^{k+1}(t_1)$ or $\mathcal{L}_G^{k+1}[0]$ can be inferred from $\lim_{t_1 \downarrow 0} \frac{d^k f(t_1)}{d^k t_1}$. This however only applies if the derivatives of z at $t_1 = 0$ and the moments of λ are finite. Now,

$$f(t_1, t_2)\Big|_{t_1, t_2=0} = z'(0)\tilde{z}'(0)h(0)h(0)E\lambda^2$$

and

$$\frac{d^k f(t_1, t_2)}{d^k t_1} \Big|_{t_1, t_2=0} = z^{k+1}(0) \tilde{z}'(0) h(0) h(0) E\lambda^2 + \dots + z(0)^{k+1} \tilde{z}'(0) h(0) h(0)^{k+1} E\lambda^{k+2}$$

Hence, by knowing all the k derivatives of $z(\cdot)$ at 0 and by knowing $E\lambda^k$, we can infer $E\lambda^{k+1}$ from $\lim_{t_1, t_2 \downarrow 0} \frac{d^k f(t_1, t_2)}{d^k t_1}$. Then, we can use $\lim_{t_1 \downarrow 0} \frac{d^k f(t_1)}{d^k t_1}$ to find the $k+1$ derivative of $z(\cdot)$ at 0. Hence we can infer all the moments of λ , which allows us to uniquely identify the distribution function G . This in turn allows us to identify $z(\cdot)$ from the distribution of t_1 . Knowing G , we can infer $h(\cdot)$ by looking at $f(t_1, t_2)$ for $\lim_{t_2 \downarrow 0}$. Then, we can identify $\tilde{z}(\cdot)$ from the distribution of t_2 . \square

This proof relies very heavily on strong assumptions about G and the smoothness of the functions h and z . It is hence not very useful empirically. In the empirical section, only the first model will therefore be considered.

3 An empirical application

The data set used derives from the Dutch Income Panel data set, which is an administrative data set based on the tax records of about 65000 randomly

picked Dutch citizens who are followed through the period 1989-1997. This period is characterised by a lack of severe economic downturns, which suggests that the assumption of a constant baseline function is not implausible for this data set (see Frijters, Lindeboom and Van den Berg, 1999). Each transition in and out of work is recorded to the day. We will focus on the young cohort starting their labour market careers in this period, i.e., those who have left school in this period, and address the hazard rates into employment as a function of present and lagged duration of non-employment. Looking only at the first 2 spells of non-employment, we are left with 2066 individuals.

First, consider the two functions

$$\int_0^{bandw} f(\widehat{t_1 = t, t_2 = s}) ds$$

and

$$\int_0^{bandw} f(\widehat{t_2 = t, t_1 = s}) ds$$

Now, $b(t) = \frac{\int_0^{bandw} f(\widehat{t_1 = t, t_2 = s}) ds}{\int_0^{bandw} f(\widehat{t_2 = t, t_1 = s}) ds}$ will approach, up to a normalisation con-

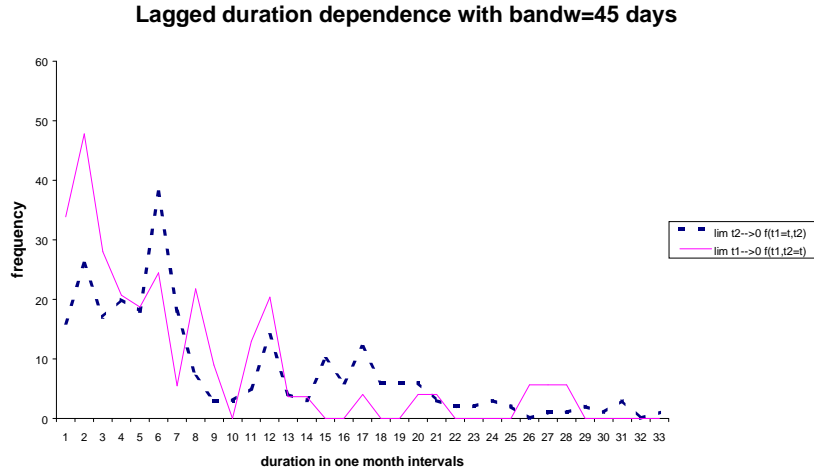


Figure 1:

stant, the lagged duration dependence function $h(t)$ for a small enough $bandw$ and a large enough number of observations under assumption 1. In Figures 1 and 2 we show both functions for two choices of $bandw$, namely 45 days and 30 days. The period of 30 days is the minimum period we need to have enough observations to look at the function $h(t)$. The results with a bandwidth of 45 days are included to see whether the shape of $h(t)$ changes quickly as we move away from a $bandw$ of zero. The functions have been re-scaled such that the means coincide.

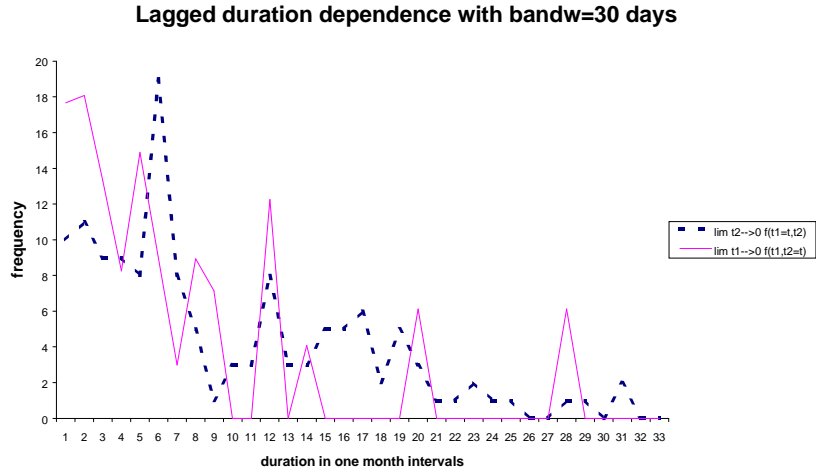


Figure 2:

Although the functions are somewhat erratic (especially with a lower bandwidth), the shape of $h(t)$ suggested by both pictures is a function increasing at first, and then declining.

Similarly, we can look at the baseline hazard function by looking at the functions $f(t_1 = t)$ and $\lim_{t_1 \downarrow 0} S(t_2 = t | t_1)$ in Figure 3 and 4, where the functions have been normalised such that the means coincide.

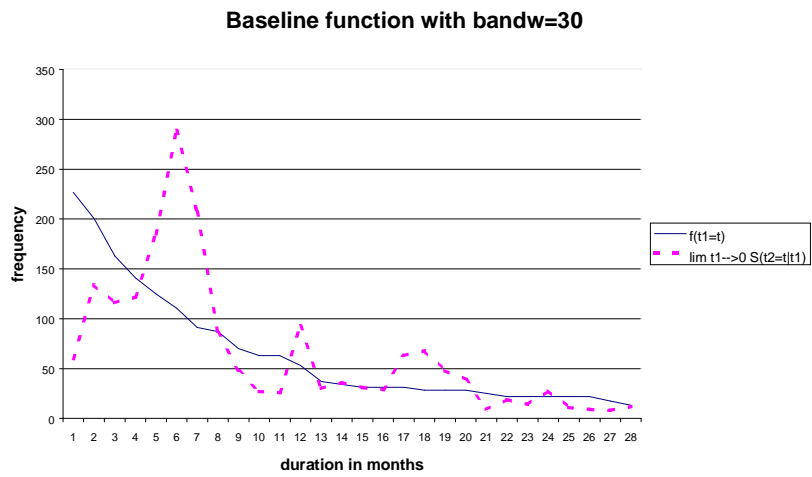


Figure 3:

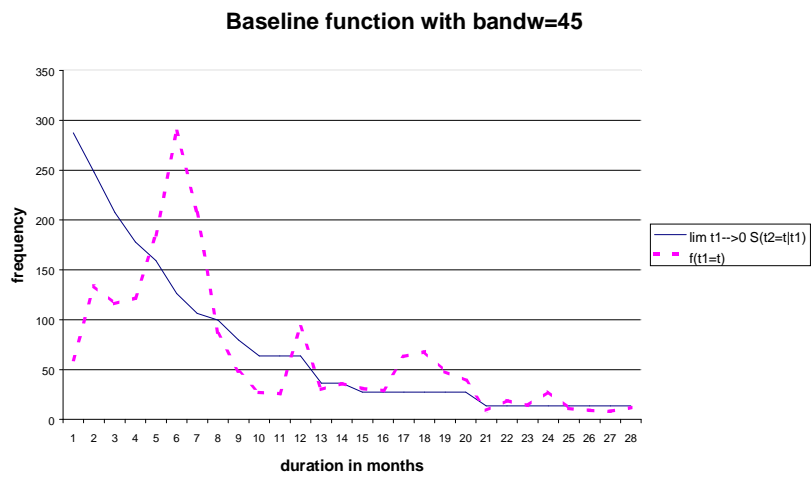


Figure 4:

which both suggest a baseline hazard which first increases sharply and thereafter decreases.

Looking somewhat closer at the lagged duration dependence function, we face the problem that due to the limited number of observations, there are points where $\int_0^{bandw} f(t_2 = \widehat{t}, t_1 = s) ds = 0$. For an estimate of $b(t)$ we can further examine, we therefore use the (Gaussian-smoothed) functions $\frac{\int_0^\infty \int_0^{bandw} e^{-(x-t)^2/w} f(t_1=\widehat{x}, t_2=s) ds dx}{\int_0^\infty e^{-(x-t)^2/w} dx}$ and $\frac{\int_0^\infty \int_0^{bandw} e^{-(x-t)^2/w} f(t_2=\widehat{t}, t_1=s) ds dx}{\int_0^\infty e^{-(x-t)^2/w} dx}$. When $w \downarrow 0$ and $bandw \downarrow 0$ and the number of observations goes to infinity such that $f(t_2, t_1)$ is approached arbitrarily well, standard limit theory arguments imply that $\hat{b}(t) = \frac{\int_0^\infty \int_0^{bandw} e^{-(x-t)^2/w} f(t_1=\widehat{x}, t_2=s) ds dx}{\int_0^\infty \int_0^{bandw} e^{-(x-t)^2/w} f(t_2=\widehat{t}, t_1=s) ds dx}$ will approach $h(t)$ in each point t as long as $\lim_{t_1 \downarrow 0} f(t_1, t_2 = t) > 0$. Because the limit distribution of $\hat{b}(t)$ could not be identified however, bootstrapping procedures will be used to obtain confidence intervals.

The function $\hat{b}(t)$ is approximated by²:

$$\hat{b}(t) = e^{\beta_0 + \beta_1 \ln(t) + \beta_2 \ln^2(t) + u_t}$$

²Because t is looked at in natural multiples of months, there are no problems with $\ln(t)$ at $t=0$.

Now, weighing each point t by the numerator of $\hat{b}(t)$ and assuming that $E[u_t|t] = 0$, the parameters β were estimated with least square techniques. Least-absolute-deviation techniques were also tried, but revealed little difference. The least-squares results were:

Table 1: Regression results of the fitted lagged duration dependence function.

	Point estimates	Bootstrapped confidence intervals from 100 bootstraps		\bar{R}^2	N*
		5%	95%		
Bandw=45*					
β_0	-2.468	-2.615	-0.305		
β_1	6.088	3.329	6.630	0.39	121
β_2	-1.978	-2.290	-1.527		
Bandw=30**					
β_0	-1.598	-2.201	0.792		
β_1	5.648	2.933	6.749	0.38	52
β_2	-2.108	-2.593	-1.813		
* Number of individuals with $t_1 < \text{bandw}$.					

From this table, we can see that $\hat{b}(t)$ increases with the length of the pre-

vious non-work spell up to about 4 months, after which it drops. The bootstrapped confidence intervals indicate that the general shape of the lagged duration dependence function (first an increase, then a decrease), is quite robust. Also, the difference between having a bandwidth of 45 or 30 days is small in the sense that the coefficients of the two cases lie well within each others confidence intervals, suggesting that the results change little as we get even closer to zero with the bandwidth.

For hazard rates into employment, these results imply that hazard rates are lower for those with longer previous spells of unemployment except for very low non-employment durations. This is somewhat in line with the findings of Omori (1997) and Trivedi and Alexander (1989), who also found significant effects of the lengths of previous non-employment. Omori interpreted his findings as evidence of a stigma effect.

4 Conclusion

Lagged duration dependence in an MPH-model without independent regressors was shown to be non-parametrically identified under mild assumptions on the baseline hazard. Estimating the lagged duration dependence function

for a Dutch administrative panel data set revealed a significant degree of lagged duration dependence: hazard rates into employment increased with the length of previous non-employment for the first 4 months of previous non-employment and decreased thereafter.

Literature

1. Blau, D.M. (1994), 'Labour force dynamics of older men', *Econometrica*, vol. 62, pp. 117-156.
2. Cross, R. (1995), *The natural rate of unemployment. Reflections on 25 years of the hypothesis*. Cambridge University Press, Cambridge.
3. Frijters, P., Lindeboom, M., Van den Berg, G. (1999), 'Herhalingswerkloosheid: persistenties op de arbeidsmarkt', report for the Ministry of Social Affairs, Staatsdrukkerij: Den Haag. *To appear shortly*.
4. Heckman, J.J. (1991), 'Identifying the hand of the past: distinguishing state dependence from heterogeneity', *American Economic Review*, 81,

pg. 71-79.

5. Heckman, J.J., Borjas, G. (1980), 'Does unemployment cause future unemployment?', *Economica*, pp. 247-283.
6. Honoré, B.E. (1993), 'Identification results for duration models with multiple spells', *Review of Economic Studies*, 60, pp. 241-246.
7. Ljungqvist, L., Sargent, T.J. (1998), 'The European unemployment dilemma', *Journal of Political Economy*, vol. 106(3), pp. 514-50.
8. Lynch, L.M. (1985), 'State dependence in youth unemployment: a lost generation?', *Journal of Econometrics*, pp. 71-84
9. Lynch, L.M. (1989), 'The youth labour market in the eighties: determinants of re-employment probabilities for young men and women', *Review of Economic Studies*, pp. 37-45.
10. Omori, Y. (1997), 'Stigma effects of nonemployment', *Economic Enquiry*, 35, pp. 394-416
11. Ridder, G. (1986), 'An event history approach to the evaluation of training, recruitment and employment programs', *Journal of Applied Econometrics*, 1, pp. 109-126.

12. Ridder, G. (1990), 'The non-parametric identification of generalised accelerated failure-time models', *Review of Economic Studies*, 57, pp. 167-82.
13. Trivedi, P.K., Alexander, J.N. (1989), 'Reemployment probability and multiple unemployment spells: a partial likelihood approach', *Journal of Business and Economic Statistics*, pp. 395-401.