DISCUSSION PAPER SERIES

IZA DP No. 898

Nonparametric Estimation of a Dependent Competing Risks Model for Unemployment Durations

Gerard J. van den Berg
A. Gijsbert C. van Lomwel
Jan C. van Ours

October 2003

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Nonparametric Estimation of a Dependent Competing Risks Model for Unemployment Durations

Gerard J. van den Berg
Free University of Amsterdam, Tinbergen Institute, CEPR and IZA Bonn

A. Gijsbert C. van Lomwel
CentER, Tilburg University

Jan C. van Ours
CentER, Tilburg University, OSA, CEPR and IZA Bonn

Discussion Paper No. 898
October 2003

IZA

P.O. Box 7240
D-53072 Bonn
Germany

Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

IZA Discussion Paper No. 898
October 2003

ABSTRACT

Nonparametric Estimation of a Dependent Competing Risks Model for Unemployment Durations*

In this paper we simultaneously analyze transitions from unemployment to employment and to nonparticipation. We estimate a dependent competing risks model with nonparametric specifications of the destination-specific duration dependence and unobserved heterogeneity terms. We use a unique population data set of French unemployment over the period 1988-1994, stratified by gender type, duration class and exit state.

JEL Classification: J64, C41

Keywords: exit rate, hazard rate, unobserved heterogeneity, duration dependence, nonparticipation

Corresponding author:

Jan C. van Ours
CentER for Economic Research
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
Tel.: +31 13 4662316
Fax: +31 13 4663042
Email: vanours@uvt.nl

* A preliminary version of this paper was distributed under the title "Individual variation in exit rates from unemployment: a nonparametric multivariate analysis using aggregate data". We thank participants at the ESEM Conference for helpful suggestions. The Département du Marché du Travail of the Ministère du Travail, de l'Emploi et de la Formation Professionelle of France kindly provided the data.

1 Introduction

Studies on unemployment durations often focus on whether the individual transition rate to employment decreases as a function of the elapsed duration. To examine this one has to correct for the presence of unobserved heterogeneity. Negative duration dependence at the individual level and unobserved heterogeneity both lead to negative duration dependence of the observed transition rate, but they have different policy implications. Negative duration dependence implies that emphasis should be put on the prevention of long-term unemployment, while unobserved heterogeneity calls for screening of newly unemployed.

To study these issues one also has to take into account that individuals may make transitions to different states than employment. Notably, they may move into nonparticipation (or "out of the labor force"). If one lumps transitions to employment and nonparticipation together, and if these transitions have different determinants, then this may lead to incorrect inference concerning the transition rate to work. Moreover, if one treats transitions to nonparticipation as right-censoring of the duration until work, and if the unobserved determinants of the transition rates to nonparticipation and work are mutually dependent, then the right-censoring is dependent, and standard duration analysis leads to incorrect inference.[1]

In this paper we estimate a model with separate transition rates to nonparticipation and work, allowing both to depend on unobserved heterogeneity terms that can be mutually dependent themselves. This is a dependent competing risks model in which the individual destination-specific transition rates (or "exit rates") have mixed proportional hazard (MPH) structures (see Lancaster, 1990, and Van den Berg, 2001, for overviews). In continuous time, such models are identified without the need to impose parametric functional form assumptions on the duration dependence terms or the bivariate distribution of the two unobserved heterogeneity terms (Heckman and Honoré, 1989, and Abbring and Van den Berg, 2003). The main identifying assumption is basically that the two exit rates should not vary with the observed covariates in exactly the same way.

We use rather unique French administrative population data over the

[1] Inference is also incorrect if one lumps the states of unemployment and nonparticipation together; Flinn and Heckman (1983) show that the transition rates to work differ between unemployment and nonparticipation.

period 1988–1994. These data are quarterly and cover individuals who are looking for full time permanent jobs. Because of the discrete nature of the data we specify and estimate a discrete time model. The main explanatory covariate is calendar time, capturing business cycle and seasonal effects. The model and the nonparametric estimation method that we develop and apply extend those in Van den Berg and Van Ours (1996) to a competing risks setting. The way in which the dependence between the unobserved heterogeneity terms is estimated provides intuition behind the identification of dependent competing risks models in general.

The outline is as follows. In the next section, the model and estimation method are introduced. We describe our data in section 3. Here, we also show how the exit probabilities are constructed from the data. Estimation results are reported in section 4. Concluding remarks are made in section 5.

2 The model and the estimation method

2.1 The Model

A spell of unemployment can end with an exit to employment (E) or to non-participation (N). We assume that at the individual level, the two conditional exit probabilities satisfy MPH-type specifications (see e.g. Van den Berg and Van Ours, 1996, and Abbring, Van den Berg and Van Ours, 2002). We only have two observed explanatory variables: the gender of the individual and calendar time. We estimate the models separately for both gender types, and therefore suppress the conditioning on the gender type. The (conditional) exit probability to destination i after t periods of unemployment, given the current value of calendar time τ, and given the unobserved heterogeneity term v_i that affects the exit to destination i, $(i = E, N)$, is now specified as

$$\theta_i(t|\tau, v_i) = \psi_{1i}(t)\, \psi_{2i}(\tau)\, v_i \tag{1}$$

The individual over-all exit probability out of unemployment is defined as $\sum_i \theta_i(t|\tau, v_i)$, so that we require the model determinants to satisfy the inequality $0 < \sum_i \theta_i(t|\tau, v_i) < 1$. The functions ψ_{1i} and ψ_{2i} represent the effect of unemployment duration and calendar time. The unobserved heterogeneity terms v_E and v_N are allowed to be correlated. We assume that the distribution G of (v_E, v_N) and the individual realizations do not change over time.

2.2 Empirical implementation

We are primarily interested in estimating the duration dependence functions and the unobserved heterogeneity distribution. As will be explained in section 3, the data provide destination-specific exit probabilities at calendar times τ and at durations t (we take t and τ to have the same measurement scale, apart from the difference in origin). These probabilities $\theta_i(t|\tau)$ are of course aggregated over unobserved heterogeneity. To express them in terms of the model determinants, we have to integrate over v_E, v_N. Consider $\theta_i(t|\tau)$ for $t = 0$, $i = E, N$;

$$\begin{aligned} \theta_i(0|\tau) &= E_{v_i}\Big[\psi_{1i}(0)\,\psi_{2i}(\tau)\,v_i\Big] \\ &= \psi_{1i}(0)\,\psi_{2i}(\tau)\,\mu_{1i} \end{aligned}$$

where,

$$\mu_{ki} := E\Big[v_i^k\Big], \quad k = 1, 2. \tag{2}$$

For $t = 1,\ i \neq j = E, N$ we get

$$\begin{aligned} &\Pr(t = 1 \wedge i|\tau) = \\ =\;& E_{v_i,v_j}\Big[\Big(1 - \psi_{1i}(0)\psi_{2i}(\tau-1)v_i - \psi_{1j}(0)\psi_{2j}(\tau-1)v_j\Big)\cdot\psi_{1i}(1)\psi_{2i}(\tau)v_i\Big] \\ =\;& \psi_{1i}(1)\psi_{2i}(\tau)\cdot E_{v_i,v_j}\Big[v_i - \psi_{1i}(0)\psi_{2i}(\tau-1)v_i^2 - \psi_{1j}(0)\psi_{2j}(\tau-1)v_iv_j\Big] \\ =\;& \psi_{1i}(1)\psi_{2i}(\tau)\cdot\Big[\mu_{1i} - \psi_{1i}(0)\psi_{2i}(\tau-1)\mu_{2i} - \psi_{1j}(0)\psi_{2j}(\tau-1)\mu_{1i1j}\Big] \end{aligned} \tag{3}$$

where,

$$\mu_{1i1j} := E\Big[v_iv_j\Big] = Cov(v_i, v_j) + \mu_{1i}\mu_{1j} \tag{4}$$

so that

$$\begin{aligned} \theta_i(1|\tau) &= \frac{\Pr(t = 1 \wedge i|\tau)}{\Pr(t \geq 1|\tau)} \\ &= \psi_{1i}(1)\psi_{2i}(\tau)\cdot\frac{\mu_{1i} - \theta_i(0|\tau-1)\frac{\mu_{2i}}{\mu_{1i}} - \theta_j(0|\tau-1)\frac{\mu_{1i1j}}{\mu_{1j}}}{1 - \theta_i(0|\tau-1) - \theta_j(0|\tau-1)} \end{aligned} \tag{5}$$

We examine ratios of observed exit probabilities. Dividing equation (5) by $\theta_i(0|\tau)$, ref. (2), we get

$$\frac{\theta_i(1|\tau)}{\theta_i(0|\tau)} = \eta_{1i} \cdot \frac{1 - \gamma_{2i}\theta_i(0|\tau-1) - \kappa_{11}\theta_j(0|\tau-1)}{1 - \theta_i(0|\tau-1) - \theta_j(0|\tau-1)} \tag{6}$$

where,

$$\eta_{1i} := \frac{\psi_{1i}(1)}{\psi_{1i}(0)}, \quad i = E, N \tag{7}$$

$$\gamma_{2i} := \frac{\mu_{2i}}{\mu_{1i}^2}, \quad i = E, N \tag{8}$$

$$\kappa_{11} := \frac{\mu_{1i1j}}{\mu_{1i}\mu_{1j}}, \quad i \neq j = E, N \tag{9}$$

If $v_N \equiv v_E$ and $\psi_{1N} \equiv \psi_{1E}$ then the model determinants of interest can be estimated from a single-risk analysis. Otherwise one has to estimate a bivariate model. We are particularly interested in the relation between v_N and v_E, because this is informative on the validity of independent right-censoring of the duration to work in a continuous-time setting.

If $v_E \perp v_N$ then $Cov(v_E, v_N) = 0$ and $\kappa_{11} = 1$. In general, the parameter κ_{11} in equation (6) is identified from the effect of the past calendar time variation in the exit to destination j on the current exit probability to destination i, i.e. on $\theta_i(1|\tau),\ i \neq j$. This fails if $\psi_{2E}(\tau) \equiv \psi_{2N}(\tau)$, i.e. if there is no independent variation in calendar time. In this case $\theta_E(0|\tau-1) \propto \theta_N(0|\tau-1)$, so only γ_{2i} + constant$\cdot\kappa_{11},\ i = E, N$ are identified from data on $t = 0, 1$. Thus, identification requires that the exit probabilities from short-term unemployment to nonparticipation and to work do not depend in exactly the same way on calendar time. This can of course be readily examined from time series of $\theta_E(0|\tau)$ and $\theta_N(0|\tau)$.

The identification condition is reminiscent of the identification condition for continuous-time models mentioned in section 1. If the covariate part of the exit probability to state i does not directly affect the individual exit probability to state $j \neq i$ but does affect the observed exit probability to state j then this indicates that there is a spurious relation between the durations by way of their unobserved determinants. The composition of the survivors at $t = 1$ then depends on the speed of the selection process for both exit destinations at $t = 0$.

For $t = 2$ one can derive similar expressions as above, leading to equations for $\theta_i(2|\tau)/\theta_i(1|\tau)$ in terms of observables $\theta_j(0|\tau-1), \theta_j(0|\tau-2), \theta_j(1|\tau-1)(j = N, E)$ and additional parameters

$$\eta_{2i} = \frac{\psi_{1i}(2)}{\psi_{1i}(1)} \quad , \; i = E, N \tag{10}$$

$$\gamma_{3i} = \frac{\mu_{3i}}{\mu_{1i}^3} \quad , \; i = E, N \tag{11}$$

$$\kappa_{12} = \frac{E\left[v_i v_j^2\right]}{\mu_{1i}\mu_{1j}^2}, \quad \kappa_{21} = \frac{E\left[v_i^2 v_j\right]}{\mu_{1i}^2\mu_{1j}}, \; i \neq j = E, N \tag{12}$$

which has the same qualitative features as (6).

We specify $\log\left(\theta_i(t|\tau)/\theta_i(t-1|\tau)\right)$ (with $t = 1, 2$) to equal the log of the corresponding model expression plus an error term. This gives a system of 4 equations. The error terms represent specification errors that are assumed to be identically $N(0, \Sigma)$ distributed over time periods. We allow the errors in the four equations to be contemporaneously related. The system is estimated by maximum likelihood. We effectively estimate 11 parameters, which are in fact values of the nonparametric function $\psi_{1i}(t)$ and moments of the nonparametrically specified G. Note that these moments do not allow for identification of the full heterogeneity distribution. Accordingly, we can only test for uncorrelatedness of the heterogeneity terms, but not for independence. Also note that the parameters describing the heterogeneity distribution appear in all four equations and are overidentified.

3 Data

We use French administrative data on individuals looking for full time permanent jobs. The data are aggregated by duration class. For the model to be applicable, the frequency at which the data are collected has to equal one over the size of the unemployment duration classes. In our empirical analysis, we use quarterly data on the exit probabilities out of the first three quarterly duration classes, over the period 1988.3-1994.4.

To derive the observed exit probabilities from the data, we use two different data sets. The first data set distinguishes unemployment by elapsed duration and gender type. These stock data enable us to calculate exit probabilities, aggregated over all possible destinations. Let $U(t|\tau)$ denote the number of unemployed in duration class t at the end of quarter τ. Then the observed exit probability equals

$$\theta(t|\tau) = \frac{U(t|\tau) - U(t+1|\tau+1)}{U(t|\tau)} \tag{13}$$

The second data set contains the number of unemployed that left unemployment in a given quarter, stratified by destination, duration of the past spell of unemployment and gender type. Five possible destinations are distinguished, viz. nonparticipation, expulsion, training, employment and others. Because of the small number of individuals in some of the categories, we group the destinations: employment and training denote the destination employment, while the other three destinations are taken together in the destination nonparticipation.

The observed destination-specific exit probabilities are now calculated in the following way. Denote the percentage of the outflow to destination i during quarter τ having been unemployed t quarters by $f_i(t|\tau)$ [2].

$$\theta_i(t|\tau) = \frac{U(t|\tau-1) - U(t+1|\tau)}{U(t|\tau-1)} f_i(t|\tau) \tag{14}$$

In order to circumvent modelling seasonal effects, we correct the raw data on the exit probabilities for seasonal effects, using the Census X11 filter (Shiskin, Young and Musgrave, 1967). It can be shown that seasonal effects in the inflow are also corrected for by this procedure. The data clearly demonstrate that the time series of the destination-specific exit probabilities are different, so that the identification condition is satisfied.

4 Estimation results

The estimation results are shown in Table 1. The exit rate to employment shows negative duration dependence during the second quarter of unemployment for both males and females. Moreover, for males there is negative duration dependence in the exit rate to nonparticipation during the first quarter

[2] An alternative approach is to divide the observed flows in the second data set by an appropriate risk set, based on the stocks in the first data set. The risk set for the flow from class t in quarter τ is of the form $\alpha U(t|\tau-1) + \beta U(t-1|\tau-1)$, with $\alpha + \beta = 1$. For reasonable values of α and β, this gives similar estimation results.

of unemployment. The other duration dependence parameters do not differ significantly from one.

The estimates of the unobserved heterogeneity distribution parameters are characterized by large standard errors. The estimates of κ_{11} indicate that the unobserved heterogeneity distributions are uncorrelated, for both gender types. If there is independence, the following restrictions should hold:

$$\kappa_{11} = 1, \;\; \kappa_{12} = \gamma_{2N}, \;\; \kappa_{21} = \gamma_{2E} \tag{15}$$

Table 1 also shows the estimation results of the model in which these restrictions are imposed. The likelihood ratio test statistic equals 0.98 for males and 2.58 for females. As the 10%-critical value of the $\chi^2(3)$ distribution is 6.25, we do not reject the independence assumption at any reasonable significance level. The parameter estimates of the duration dependence parameters are close to the estimates in the unrestricted model. Duration dependence in the exit rate to employment is similar for males and females, with significant negative duration dependence during the second quarter of unemployment.

Concerning the unobserved heterogeneity parameters, note that a γ_{2i} estimate significantly larger than one implies that $var(v_i) > 0$, i.e. that unobserved heterogeneity is present. For both males and females, we find unobserved heterogeneity in the exit rate to both employment and nonparticipation, which is hardly surprising, given the high level of aggregation in the data. Further, $\gamma_{3i} - \gamma_{2i}^2$ does not differ significantly from zero, for each gender. From Shohat and Tamarkin (1970), this implies that the marginal distributions of v_i can be accurately described by a discrete distribution with one positive point of support and one point of support equal to zero. This is convenient for the interpretation of the results, for two reasons. First, for bivariate discrete distributions with two by two points of support, uncorrelatedness is equivalent to independence. Secondly, it is easy to summarize the bivariate distribution in words. There is not much difference across gender. About two third has a high probability of finding a job, while one third never finds job. The probability of becoming nonparticipant is positive for almost all unemployed. This makes sense, given the fact that one can personally decide to become nonparticipant.

5 Conclusion

The individual conditional exit probabilities to employment and nonparticipation are uncorrelated across individuals, for males as well as females. This suggests that one may treat exits to nonparticipation as independent right-censoring of the duration until work, at least under the current set of covariates. These results do not depend on arbitrary functional-form assumptions concerning duration dependence or the unobserved heterogeneity distribution. Moreover, we find that the duration dependence patterns and heterogeneity distributions of the conditional exit probabilities to employment and nonparticipation are different from each other. The exit to employment displays negative duration dependence after one quarter of unemployment, while the exit to nonparticipation does not display duration dependence for females, and negative duration dependence during the first quarter for males. Therefore, in the outflow from unemployment, it is important to distinguish between exits to employment and exits to nonparticipation. The differences in dynamics between the two transitions have to be taken into account.

References

Abbring, J.H., G.J. van den Berg and J.C. van Ours (2002), "The anatomy of unemployment dynamics", *European Economic Review*, 46, 1785–1824.

Abbring, J.H. and G.J. van den Berg (2003), "The identifiability of the mixed proportional hazards competing risks model", *Journal of the Royal Statistical Society Series B*, 65, 701–710.

Flinn, C.J. and J.J. Heckman (1983), "Are unemployment and out of the labor force behaviorally distinct labor force states", *Journal of Labor Economics*, 1, 28–42.

Heckman, J.J. and B.E. Honoré (1989), "The identifiability of the competing risks model", *Biometrika*, 76, 325–330.

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

Shiskin, J., A.H. Young and J.C. Musgrave (1967), "The X-11 variant of the census method II seasonal adjustment program", Technical paper 15, Bureau of the Census, U.S. Department of Commerce.

Shohat, J.A. and J.D. Tamarkin (1970), *The problem of moments* , American Mathematical Society, Providence.

Van den Berg, G.J. and J.C. van Ours (1996), "Unemployment dynamics and duration dependence", *Journal of Labor Economics*, 14, 100–125.

Van den Berg, G.J. (2001), "Duration models: Specification, identification, and multiple durations", in J.J. Heckman and E. Leamer, editors, *Handbook of Econometrics, Volume V*, North Holland, Amsterdam.

Table 1. Parameter estimates
(standard errors in parentheses)

	Unrestricted model				Restricted model			
	Males		Females		Males		Females	
unobserved heterogeneity distribution								
γ_{2E}	1.372	(0.137)	1.324	(0.148)	1.446	(0.070)	1.519	(0.098)
γ_{3E}	1.002	(0.079)	1.005	(0.167)	2.030	(0.262)	2.170	(0.469)
γ_{2N}	1.254	(0.314)	1.215	(0.242)	1.432	(0.129)	1.570	(0.127)
γ_{3N}	1.005	(0.189)	1.007	(0.335)	1.430	(0.650)	1.793	(0.867)
κ_{11}	1.120	(0.192)	1.284	(0.167)				
κ_{12}	0.939	(0.572)	1.368	(0.702)				
κ_{21}	2.372	(0.429)	2.423	(0.574)				
duration dependence								
η_{1E}	1.053	(0.037)	1.058	(0.032)	1.041	(0.028)	1.047	(0.032)
η_{2E}	0.929	(0.031)	0.938	(0.028)	0.928	(0.023)	0.932	(0.028)
η_{1N}	0.907	(0.041)	0.985	(0.030)	0.918	(0.037)	0.986	(0.030)
η_{2N}	0.991	(0.048)	1.024	(0.037)	0.990	(0.038)	1.028	(0.037)
log-likelihood								
	331.71		351.49		331.22		350.20	

IZA Discussion Papers

No.	Author(s)	Title	Area	Date
884	J. Masso A. Heshmati	The Optimality and Overuse of Labour in Estonian Manufacturing Enterprises	4	10/03
885	A. Constant K. F. Zimmermann	The Dynamics of Repeat Migration: A Markov Chain Analysis	1	10/03
886	J. J. Dolado M. Jansen J. F. Jimeno	On-the-Job Search in a Matching Model with Heterogenous Jobs and Workers	1	10/03
887	B. Irlenbusch D. Sliwka	Transparency and Reciprocal Behavior in Employment Relations	7	10/03
888	W. Koeniger	Collective Dismissal Cost, Product Market Competition and Innovation	3	10/03
889	D. E. Wildasin	Fiscal Policy, Human Capital, and Canada-US Labor Market Integration	2	10/03
890	M. Bratti L. Mancini	Differences in Early Occupational Earnings of UK Male Graduates by Degree Subject: Evidence from the 1980-1993 USR	6	10/03
891	L. Flood E. Pylkkänen R. Wahlberg	From Welfare to Work: Evaluating a Proposed Tax and Benefit Reform Targeted at Single Mothers in Sweden	6	10/03
892	B. T. Hirsch	What Do Unions Do for Economic Performance?	5	10/03
893	K. Sabirianova Peter	Skill-Biased Transition: The Role of Markets, Institutions, and Technological Change	4	10/03
894	R. Winkelmann	Parental Separation and Well-Being of Youths	7	10/03
895	J. M. Fitzgerald D. C. Ribar	Transitions in Welfare Participation and Female Headship	3	10/03
896	S. W. Polachek	What Can We Learn About the Decline in U.S. Union Membership from International Data?	2	10/03
897	M. Brown A. Falk E. Fehr	Relational Contracts and the Nature of Market Interactions	7	10/03
898	G. J. van den Berg A. G. C. van Lomwel J. C. van Ours	Nonparametric Estimation of a Dependent Competing Risks Model for Unemployment Durations	1	10/03

An updated list of IZA Discussion Papers is available on the center's homepage www.iza.org.