REGERINGSKANSLIET

Promemoria

2000-03-01

Ministry of Finance

Economic Affairs Department
Peter Ericson
Tel: + 46 8 405 1591

Unemployment in Sesim

This is a part of the documentation of Sesim.

Introduction

An individual is defined to be unemployed a given year if the contribution of unemployment insurances is a major part of the indivdual's income. At risk to unemployment are indivduals that enter the labor market in Sesim. That is, all individuals who are not classified as children, pensioners, students or mothers with a newborn. A dynamic discrete paneldata model is then used to assign a probability of unemployment to all individuals at risk. Whether an individual actually is predicted to become unemployed a given year depends on what number the assigned probability is compared to. With no alignment all individuals with a probability greater than 0.5 are predicted to become unemployed the actual year.

The model

Let y_{it} take the value one if individual i is unemployed year t, and zero otherwise. A latent variabel y_{it}^* is then used to explain the observed outcome y_{it} by the rule,

$$\begin{aligned} y^*_{it} &= \beta' X_{it} + \lambda y_{it-1} + \alpha_i + u_{it} \\ y_{it} &= 1 \;\; if \; y_{it}^* > 0 \\ y_{it} &= 0 \;\; else \end{aligned} \qquad (1)$$

This is one of the models discussed by Heckman (1981). Compared to the ordinary probit or logit model the lagged observed outcome (y_{it-1}) and the parameter α_i cause some estimation problems. In this application the parameters in model (1) are estimated by maximum-likelihood (ML), under the assumption that both α_i and u_{it} are normally distribution. The initial-condition problem is "solved" by assuming that y_{i0} is random variables with a probability distribution $P(y_{i0}=1)=F(Z_{i0}'\delta)$.

Data

The sample used in the empirical analysis is extracted from the Swedish panel data LINDA, which provides information on a sample of about 300,000 housholds. After some exclusions and the enforcement of the criteria that all indivduals should be observed during the period 1990-1998 we end up with a sample of 127,000 indivduals.

The observed exogenous variables used to explain the probability of being unemploymed in 1990, that is t=0, is presented in Table 1., while the variables used to explain the same probability between 1991-1998 is presented in Table 2.

Table 1. Variables used to explain unemployment in 1990

Variable	*Description*
age	years of age
age2	age^2/40
secondary	binary variable taking the value one if secondary education is the highest education
tertiary	binary variable taking the value one if tertiary education is the highest education
graduated	binary variable taking the value one if the person is graduated
sex	binary variable taking the value one if the person is a female
sex*secondary	interaction
sex*graduated	interaction

Table 2. Variables used to explain unemployment between 1991-1998

Variable	*Description*
age24	binary variable indicating whether a person is below 25 years of age
age25_44	binary variable indicating whether a person is between 25-44 years of age
secondary	binary variable taking the value one if secondary education is the highest education
tertiary	binary variable taking the value one if tertiary education is the highest education
graduated	binary variable taking the value one if the person is graduated
sex	binary variable taking the value one if the person is a female
unemployment	overall unemployment rate (percent)

Results

The estimated paramaters are presented in Table 3. and Table 4.

Table 3. Estmation results 1990

Parameters	*estimates*	*standard errors*
intecept	-2.45	0.020
age	0.0154	0.00161
age2	-0.00486	0.000541
secondary	-0.212	0.00836
tertiary	-0.328	0.00427
graduated	-0.503	0.00436
sex	0.223	0.0547
sex*secondary	-0.0661	0.0169
sex*graduated	-0.520	0.110

Table 4. Estimation results 1991-1998

Variable	*estimates*	*standard errors*
intercept	-2.46	0.0168
y_{t-1}	1.36	0.0100
age24	0.671	0.0120
age25_44	0.218	0.00532
secondary	-0.161	0.0107
tertiary	-0.430	0.0114
graduated	-0.581	0.0136
sex	-0.117	0.00978
unemployment	0.0698	0.00161
std(α_i)	0.652	0.00364

The results are in line with the expectations about how the probabilty to become unemployed is affected by individual characteristics. Investment in human capital and experience (years of age) both decrease the risk. An overall high unemployment does of course increase the individual probability of beeing unemployed. The inclusion of this variable is a simple way to get an interaction between prognoses on a macro-level and the individual risks on a micro-level. A person who was unemployed the last period have a much higher risk to keep that status compared to persons doing something else the last period.

To get some variation between the variables used in 1990 and the variables used between 1991-1998 two interactions are added in 1990. Also are the variabels catching the age effect modelled in a slightly different way. Instead of using binary variables to identify different agegroups, the variable ”years of age” and its second moment are utilized. A conclusion is that female generally have a lower probability to become unemployed, but, at least in 1990 this is only true if they also have a secondary eduaction, or are graduated. (The interaction between

sex and tertiary education are not included in the model since it was not significant in a preestimation to get good starting values).

References

Heckman J.J., 1981. Statistical models for discrete panel data, in : Manski, C.F, McFadden, D., (Eds.) Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge, MA, pp. 114-178.