Shooting at Moving Targets: Short versus Long Term Effects of Anti-Poverty Policies

R. De Blander and I. Nicaise
HIVA - K.U.Leuven

4th November 2005

Abstract

In this paper we model employment and poverty states as a discrete choice first-order Markov process, taking into account the endogeneity of schooling and of the initial states. Using this model, we asses the impact of different policies on the poverty dynamics by ex-post microsimulation. Each policy evaluated in this paper can be considered an exemplar of one of the competing paradigms guiding social policy: the traditional welfare state, the active welfare state or the knowledge-based society.

JEL classification: C35, I28, I32, I38, J62, J68
Keywords: Poverty dynamics, Unemployment, Markov models

Introduction

It is a well established fact that the mobility into and out of poverty is rather high. Bane and Elwood (1986), for instance, report that "60 percent of those persons just beginning a spell of poverty will exit within two years". Some of the consequences of this observation are that the fraction of persistent poor is smaller and the fraction of people that were ever poor is larger than the instantaneous fraction of poor. Apart from obvious implications for policy, similar observations, combined with increased data availability, have lead to a remarkable output focusing on the dynamics of social exclusion and poverty.

Among the recent research on poverty dynamics we find relatively simple models for poverty entry and exit with varying degrees of attention for different complications. Cappellari and Jenkins (2002), for instance, use a first-order Markov chain model and control for endogeneity of the initial conditions and for attrition. Breen and Moisio (2004) use a latent mover-stayer model with correction for measurement error in the poverty status. A more sophisticated approach models the duration before exit out of and re-entry into poverty (Callens (2004) for 10 EC countries, Canto (2002) for Spain, Devicienti (2001) for the UK, Finnie and Sweetman (2003) for Canada, Stevens (1999) for a US sample, taking into account unobserved heterogeneity). Income mobility has also been modelled in a similar fashion (Cappellari (2001), DiPrete and McManus (2000), Jenkins (2000), Cantó (2000), Böheim et al. (1999), Stewart and Swaffield (1999)). A different but related research question considers persistence in welfare benefit uptake (Gustafsson et al. (2004), see also Alcock (2004) for a general discussion and Noble et al. (1998) for an overview).

In this paper we present a joint Markov model for employment and poverty, conditional on educational attainment. This model is estimated in several stages. The three observed schooling levels are modelled using an ordered logit model. We consider the states employment and unemployment and estimate them with a state-dependent logit model. With respect to poverty we discern three states: insufficient protection (IP), minimum income (MI) and non-poverty (NP), which are state-dependently estimated using a multinomial logit model. To correct for the endogeneity of schooling and employment we include the generalized residuals (Cox and Snell (1968), Gouriéroux et al. (1987)) of the previous stages in each regression. This amounts to Heckman's (1976, 1978) control function approach adapted for ordered and multinomial choice equations (Dubin and McFadden (1984)). To correct for the initial selection effect, we also include control functions generated from static labour and poverty equations for the initial period. As a consequence of the joint modelling of poverty and social assistance benefit uptake, the poverty line we consider is the official threshold for obtaining income support. Finally, we resort to ex-post microsimulation (Merz (1991)) of three basic strategies against poverty: increasing the coverage of the minimum income, activation of the unemployed poor and raising the educational level of vulnerable groups. Using the estimated Markov model, we simulate the impact of these anti-poverty policies for the representatives of the respective target groups present in our sample over the period they were observed, i.e. a time horizon of five years. In the next section, we present our econometric model. Section two deals with the data and estimation results. In section three the simulation results are presented. Finally, section four concludes.

1 Empirical Model

We assume that schooling is determined by the latent propensity for education

$$s_i^* = \alpha' x_i + u_i, \tag{1}$$

and we observe

$$\begin{aligned} s_i &= 1, && \text{if } s_i^* \leq \mu_1 \\ &= 2, && \text{if } \mu_1 < s_i^* \leq \mu_2 \\ &= 3, && \text{if } \mu_2 < s_i^*. \end{aligned}$$

Assuming that u_i are IID according to a type I extreme-value distribution, the probabilities to observe $s_i = 1, 2, 3$ are given by

$$\begin{aligned}
\Pr[s_i = 1 \mid x_i] &= \Lambda(\mu_1 - \alpha' x_i) \\
\Pr[s_i = 2 \mid x_i] &= \Lambda(\mu_2 - \alpha' x_i) - \Lambda(\mu_1 - \alpha' x_i) \\
\Pr[s_i = 3 \mid x_i] &= 1 - \Lambda(\mu_2 - \alpha' x_i), \qquad (2)
\end{aligned}$$

with $\Lambda(y) = \frac{e^y}{1+e^y}$.

In the first period, individual i is at work, when her employability

$$w_{i1}^* = \beta_1' y_{i1} + \beta_2 d_{s_2;i} + \beta_3 d_{s_3;i} + v_{i1} \qquad (3)$$

is higher than some threshold (0). In that case we observe $w_i = 1$, otherwise $w_i = 0$. We allow the errors u_i and v_{it} to be correlated, but assume a linear dependency between both error terms

$$v_{i1} = \beta_4 u_i + \eta_{i1}. \qquad (4)$$

Taking the expectation of (3) with respect to x_{it}, y_{it} and s_{it}, results in

$$\mathrm{E}[w_{i1}^* \mid x_i, y_{i1}, s_i] = \beta_1' y_{i1} + \beta_2 d_{s_2;i} + \beta_3 d_{s_3;i} + \beta_4 \mathrm{E}[u_i \mid x_i, y_{i1}, s_i].$$

When the instruments for s_i are valid, it holds that $\mathrm{E}[u_i \mid x_i, y_{i1}, s_i] = \mathrm{E}[u_i \mid x_i, s_i]$. This last expression is also termed the generalized residual (Cox and Snell (1968), Gouriéroux et al. (1987)) of the schooling model (1-2). Its inclusion in equation (3) removes the bias due to the correlation between u_{it} and v_{it}. This can also be considered an extension of Heckman's (1976, 1978) control function approach to an ordered logistic choice equation, for which the expressions are given in the appendix.

From the second period on, the individual's employability is assumed to depend on the previous state, and its expectation can be written as

$$\begin{aligned}
\mathrm{E}[w_{it}^* \mid x_i, s_i, y_{i1}, w_{i1}, y_{it}, w_{i;t-1}] = \; & \beta_{1;w_{i;t-1}}' y_{i1} + \beta_{2;w_{i;t-1}} d_{s_2;i} + \beta_{3;w_{i;t-1}} d_{s_3;i} \\
& + \beta_{4;w_{i;t-1}} \tilde{u}_i + \beta_{5;w_{i;t-1}} \tilde{v}_{i1},
\end{aligned}$$

under the following linearity assumption

$$v_{it;w_{i;t-1}} = \beta_{4;w_{i;t-1}} u_i + \beta_{5;w_{i;t-1}} v_{i1} + \eta_{it}, \qquad (5)$$

and with $\tilde{u}_i = \mathrm{E}[u_i \mid x_i, s_i]$ and $\tilde{v}_{i1} = \mathrm{E}[v_{i1} \mid y_{i1}, s_i, w_{i1}]$.

The same reasoning can again be applied for the poverty equations, with the modification that the endogeneity of the initial poverty state will be controlled for by two control functions. For the first period, the propensity for individual i of being in state IP or MI is given by

$$
\begin{aligned}
p^*_{i1;1} &= \gamma'_{1;1} z_{i1} + \gamma_{2;1} d_{s_2;i} + \gamma_{3;1} d_{s_3;i} + \gamma'_{4;1} w_{i1} + \gamma_{5;1} u_i + \gamma_{6;1} v_{i1} + \varepsilon_{i1;1} \\
p^*_{i1;2} &= \gamma'_{1;2} z_{i1} + \gamma_{2;2} d_{s_2;i} + \gamma_{3;2} d_{s_3;i} + \gamma'_{4;2} w_{i1} + \gamma_{5;2} u_i + \gamma_{6;2} v_{i1} + \varepsilon_{i1;2},
\end{aligned} \tag{6}
$$

while for the other periods, they depend on the previous state and we need to control for the initial conditions

$$
\begin{aligned}
\mathrm{E}\left[p^*_{it;1;p_{i,t-1}}\right] &= \gamma'_{1;1;p_{i,t-1}} z_{it} + \gamma_{2;1;p_{i,t-1}} d_{s_2;i} + \gamma_{3;1;p_{i,t-1}} d_{s_3;i} + \gamma'_{4;1;p_{i,t-1}} w_{it} \\
&\quad + \gamma_{5;1;p_{i,t-1}} \tilde{u}_i + \gamma_{6;1;p_{i,t-1}} \tilde{v}_{it} + \gamma_{7;1;p_{i,t-1}} \tilde{\varepsilon}_{i1;1} + \gamma_{8;1;p_{i,t-1}} \tilde{\varepsilon}_{i1;2} \\
\mathrm{E}\left[p^*_{it;2;p_{i,t-1}}\right] &= \gamma'_{1;2;p_{i,t-1}} z_{it} + \gamma_{2;2;p_{i,t-1}} d_{s_2;i} + \gamma_{3;2;p_{i,t-1}} d_{s_3;i} + \gamma'_{4;2;p_{i,t-1}} w_{it} \\
&\quad + \gamma_{5;2;p_{i,t-1}} \tilde{u}_i + \gamma_{6;2;p_{i,t-1}} \tilde{v}_{it} + \gamma_{7;2;p_{i,t-1}} \tilde{\varepsilon}_{i1;1} + \gamma_{8;1;p_{i,t-1}} \tilde{\varepsilon}_{i1;2},
\end{aligned} \tag{7}
$$

where $\tilde{\varepsilon}_{i1;1}$ and $\tilde{\varepsilon}_{i1;2}$ are again given in the appendix.

Two further remarks are in order. First, we do not fully exploit the panel structure of our data, nor of the fact that several members of the same household may be included in the sample. However, White (1982) guarantees that, when the parameters are consistently estimated, their variance is given by

$$
\mathrm{E}\left[\frac{\partial^2 LL}{\partial\theta\partial\theta'}\right]^{-1} \mathrm{E}\left[\frac{\partial LL}{\partial\theta} \cdot \frac{\partial LL}{\partial\theta'}\right] \mathrm{E}\left[\frac{\partial^2 LL}{\partial\theta\partial\theta'}\right]^{-1},
$$

a quantity which can easily be estimated. We apply this method since the use of a fixed effect estimator would not allow identification of the effect of schooling, and random effects merely allows some efficiency gains. Second, the fact that the control function needs to be estimated in a previous stage, is taken care of by the δ-method. For more details, see the appendix.

2 Estimation

2.1 Data

We consider three educational levels: primary or lower secondary education, upper secondary education and higher education; two labour market states: part- or full-time working (W) and non-working (U); and three poverty states:

- Insufficient protection (IP): family income lies below the legally guaranteed minimum income. For some reason, people in this state forego income support[1]. Note that some

[1] For a model of non-uptake see Riphahn (2001).

individuals in this state may draw an income from work or (other) social benefits. However this does not lift them above the official poverty line.

- Social assistance or minimum income (MI): the municipal social service pays the difference between earned income and the guaranteed minimum income.
- Non-poverty (NP): family income lies above the minimum income, irrespective whether it consists of wages or social security benefits.

From the above distinction it is clear that our definition of the poverty threshold is identical to the Belgian government's cut-off point for receiving social assistance. In the literature this poverty line is considered to result in an underestimation of the number of poor people. We nevertheless maintain it for the following reasons. First this threshold distinguishes a qualitatively different part of the population, those entitled to income support. Second, a higher poverty line would blur the difference between the IP and the MI states, making a model accounting for both poverty and social assistance dynamics much more difficult.

The dataset we use is a subset of the Panel Study of Belgian Households (Mortelmans et al. (2004)), from which we retained all individuals out of school but not yet (early) retired, since pensioners, children and students are excluded from certain states. Our sample thus consists of 5380 individuals, with monthly observations for on income and labour market status during the period 1993-'97.[2]

2.2 Results

2.2.1 Schooling

The first equation estimates the probability for an individual to have achieved a certain level of education, using an ordered logit model. The results of this estimation procedure are given in Table 1, from which the following conclusions can be drawn. Women, catholics and younger birth cohorts generally have achieved higher educational levels. Parental social status also has the expected positive effect on the educational performance of their offspring. The influence of nationality at birth is statistically negligible after controlling for the other determinants. Likelihood-ratio tests for each group of dummies describing the same underlying continuous variable, are given in Table 2. Mother's education is statistically insignificant at 1%, but significant at 5%, while the other determinants are all significant at 1%.

[2]We reconstructed monthly income data by combining the yearly income and monthly activity variables from the panel. For a detailed account of the methodology, see Nicaise et al. (2004).

Table 1: Determinants of educational attainment

Variable	b	$s.e.$	p
Gender	0.0924	0.0542	0.088
Catholic	0.1848	0.0743	0.013
Born in Belgium	0.0445	0.0943	0.637
Born 40s	0.6921	0.1073	0.000
Born 50s	0.8175	0.1009	0.000
Born 60s	1.1071	0.1034	0.000
Born 70s	0.8418	0.1238	0.000
Father unemployed	-0.0811	0.3049	0.790
Father blue-collar worker	-0.5581	0.0892	0.000
Father white-collar worker	0.4188	0.1094	0.000
Father self-employed	0.0410	0.1124	0.715
Father executive	0.4668	0.1193	0.000
Mother unemployed	-0.1891	0.0929	0.042
Mother blue-collar worker	-0.3895	0.1325	0.003
Mother white-collar worker	-0.2036	0.1337	0.128
Mother self-employed	-0.3321	0.1292	0.010
Mother executive	-0.3941	0.2631	0.134
Father no education	-0.1508	0.1473	0.306
Father lower sec. educ. or less	0.2329	0.1111	0.036
Father higher sec. educ.	0.7412	0.1231	0.000
Father higher education	1.2373	0.1340	0.000
Mother no education	-0.6437	0.1414	0.000
Mother lower sec. educ. or less	0.1195	0.1116	0.284
Mother higher sec. educ.	0.6733	0.1278	0.000
Mother higher education	0.8638	0.1466	0.000
intercept 1	0.4262	0.1576	
intercept 2	2.0818	0.1608	
# individuals	5380		
χ^2_{25}	1348.53		0.000

Table 2: Likelihood ratio tests for each group of socio-economic background dummies in table 1

	χ^2_{df}	$d.f.$	p
Cohort	123.4462	4	0.000
Employment father	139.5040	5	0.000
Employment mother	12.1002	5	0.033
Education father	136.5826	4	0.000
Education mother	120.5392	4	0.000

2.2.2 Employment

The work status dummy-variable (1: working / 0: non-working) was estimated dynamically[3] from the second period onward, depending on the work status in the previous period, using a logit model. This dynamic model will subsequently be used to simulate the effects of activation and education policies on the poverty transitions. A priori we presume that those policies will have persistent effects both on labour market and poverty dynamics. In Table 3 we see that most regressors behave as expected. Education boosts both the probability to get and to stay at work, with the effect of higher education on access to work being almost twice as high as its effect on non-exit. Younger people have a higher probability of access to work, but also of job loss. In other words, youth unemployment is more volatile, while non-employment at later ages is more persistent. Unemployment in the countryside is also more persistent.

Women and single parents with more children and persons living in large families or in bad health, all have a lower probability of getting or staying at work. The lower probability of being at work experienced by non-Belgian EU citizens stems mainly from their slightly lower probability of keeping their job. The lower rate of employment of non-Europeans, on the other hand, is mainly caused by a lower probability of access to work. We nevertheless think that labour market policies targeted at ethnic minorities should focus on inequalities in job retention as well as on discrimination in hiring.

Finally, the control function derived from the education equation has no significant effect on labour market transitions, but the correction terms with respect to initial employment status are significant. The latter effect points to the presence of an individual-specific error component. Indeed, the control functions for the initial conditions can be considered as a measurement (with error) of the individual-specific error component. There is thus a strong selection effect in the labour market dynamics. On top of this selection effect, there is persistence in the probability of being at work: a likelihood-ratio test strongly rejects the null hypothesis of equality of the coefficients in both job entry and exit probabilities, i.e. the null of a static model.

2.2.3 Poverty

In this section we discuss the probability of finding oneself in one of the three poverty states, conditional on the state in the previous period[4]. The poverty state is given by a trinomial nominal variable, with reference category NP.

A first striking conclusion with respect to entry into poverty, in Table 4, is the insignificant influence of employment status on the probability to enter MI. However, being at work significantly diminishes the probability of becoming insufficiently protected. Also, the self-employed

[3] The static probability of working in the initial period is not reported here (see De Blander and Nicaise (2005) for details and discussion).

[4] Again we refer to De Blander and Nicaise (2005) for the static estimation for the initial period.

Table 3: Determinants of employment

previous state	unemployed			employed		
Variable	b	$s.e.$	p	b	$s.e.$	p
Higher sec. educ.	0.5066	0.0992	0.000	0.4583	0.0987	0.000
Higher education	1.3334	0.1573	0.000	0.7014	0.1500	0.000
Age $<$ 25 y.	3.7765	0.2085	0.000	-0.3954	0.1461	0.007
Age 25-34 y.	3.8249	0.2005	0.000	1.0907	0.1403	0.000
Age 35-44 y.	3.4096	0.2012	0.000	1.6700	0.1405	0.000
Age 45-54 y.	2.2956	0.2076	0.000	1.4916	0.1431	0.000
Gender	-0.7343	0.0681	0.000	-0.9813	0.0573	0.000
Cohabiting	0.2052	0.0749	0.006	0.4328	0.0621	0.000
Household size	-0.0373	0.0348	0.285	-0.1398	0.0281	0.000
# kids $<$ 12 y.	-0.2113	0.0462	0.000	0.0095	0.0393	0.809
# kids 12-16 y.	-0.2814	0.0885	0.001	-0.1530	0.0754	0.042
Poor health	-0.2415	0.0367	0.000	-0.3118	0.0382	0.000
EU citizen	-0.2684	0.1416	0.058	-0.3493	0.1345	0.009
non-EU citizen	-1.1684	0.2156	0.000	-0.4311	0.1867	0.021
Urban residence	0.1481	0.0636	0.020	-0.1611	0.0564	0.004
Brussels region	-0.2369	0.1114	0.033	-0.0880	0.0901	0.329
Walloon region	-0.2178	0.0625	0.000	-0.2538	0.0568	0.000
Ec. growth	-0.0590	0.0363	0.104	0.0258	0.0353	0.465
Unempl. rate	0.1548	0.1591	0.330	-0.1737	0.1390	0.211
CF* schooling	-0.0232	0.0398	0.561	-0.0305	0.0394	0.439
CF initial cond.	0.3870	0.0484	0.000	0.4910	0.0613	0.000
Constant	-7.0087	1.4495	0.000	6.0135	1.2635	0.000
# observations	69442			152933		
χ^2_{21}	2225.46		0.000	2928.97		0.000

*CF = control function.

are more likely to enter IP compared to MI.

Both probabilities are inversely proportional to educational level, and decline with age[5]. Note that the age-class 45-54 does not fit into this pattern. Younger people are especially vulnerable to becoming poor, a finding which calls for special attention for this target group. Women, singles and smaller families or families with children have a higher probability of becoming poor. City dwellers and people in bad health have a higher likelihood of obtaining social assistance, as do people living in the Brussels area or Wallonia.

The control functions for the endogeneity of schooling and the initial working conditions are not significant. The initial poverty conditions and the work transitions are endogenous. People who find (or keep) a job against the (observed) odds, are also more likely to stay out of poverty.

Table 5 reports the influence of determinants on the probability of leaving the IP state. A job significantly diminishes the likelihood of a prolonged stay in IP or a transition to MI. This likelihood is also inversely proportional to educational level. The IP state tends to be relatively more persistent for the self-employed. Cohabitation and household size improve the chances of exit from poverty, but children decrease this probability. Bad health seems to "help" people in finding ass istance.

The initial conditions of both work and poverty do not seem to influence the exit probability from IP. People with a high unobserved component in their education also tend to stay longer in IP, while people with a high unobserved component in their transitions into work have a higher exit probability from poverty.

We finally remark that our sample contains only 24 transitions from IP to MI, which corresponds to a transition probability of 0.53%.

In Table 6 we describe the transition probabilities from the MI state into IP or MI. As before, the number of transitions from MI to IP is very low and corresponds to a transition probability of only 0.71%. An obvious conclusion is that the mode of poverty in which people are situated is very persistent.

Having a job or a better diploma and living together promote exits from social assistance, while household size and the number of adolescent children increases MI persistence. The control functions for education and the initial work conditions are not significant. On the other hand, the initial poverty conditions as well as the contemporaneous work transitions are significantly endogenous. As with exit from IP, a high unobserved component in the work transitions results in a higher exit probability from MI.

Our model generates some insights into direct and indirect causal connections between risk factors and outcomes in terms of poverty:

- Socio-economic background (in terms of parental education and occupation) strongly influences a person's educational achievement and hence his risk of becoming and staying

[5] Age-class 55-64 is our reference category.

Table 4: Probability of becoming poor

From non-poverty to	Insufficient Protection			Social Assistance		
Variable	b	$s.e.$	p	b	$s.e.$	p
Working	-1.1015	0.1796	0.000	-0.3233	0.3368	0.337
Higher sec. educ.	-0.2749	0.1815	0.130	-1.4547	0.4490	0.001
Higher education	-0.6826	0.3094	0.027	-2.6427	0.7155	0.000
Self-employed	1.4407	0.1676	0.000	-1.4834	1.0328	0.151
Age $<$ 25 y.	0.9936	0.2503	0.000	1.7335	0.6017	0.004
Age 25-34 y.	0.7677	0.2239	0.001	0.9862	0.5460	0.071
Age 35-44 y.	0.2994	0.2156	0.165	0.8589	0.5077	0.091
Age 45-54 y.	0.3053	0.2008	0.128	1.1961	0.4683	0.011
Gender	0.1951	0.1118	0.081	0.3111	0.2314	0.179
Cohabiting	-0.5229	0.1172	0.000	-0.8804	0.2578	0.001
Household size	-0.4502	0.0759	0.000	-0.1549	0.1131	0.171
# kids $<$ 12 y.	0.2260	0.0924	0.014	0.2131	0.1431	0.137
# kids 12-16 y.	0.5921	0.1487	0.000	0.5276	0.2393	0.027
Bad health	0.0751	0.0626	0.230	0.2953	0.1095	0.007
EU citizen	0.1414	0.2298	0.538	0.3154	0.3937	0.423
non-EU citizen	0.2084	0.2712	0.442	0.5502	0.4136	0.183
City	-0.1556	0.1044	0.136	0.6734	0.2233	0.003
Brussels region	-0.6159	0.1858	0.001	0.8051	0.3176	0.011
Walloon region	0.0072	0.1055	0.945	0.5811	0.2535	0.022
Ec. growth	0.0326	0.0433	0.452	0.1496	0.0945	0.113
Unempl. rate	0.4513	0.1760	0.010	0.4836	0.3708	0.192
CF* schooling	0.0986	0.0813	0.225	0.1791	0.1865	0.337
CF init. cond. work	-0.0869	0.0556	0.118	-0.0616	0.1100	0.576
CF work	-0.4955	0.0419	0.000	-0.2617	0.1816	0.150
CF init. cond. Insuff. prot.	0.3944	0.0441	0.000	0.0441	0.1150	0.702
CF init. cond. Min. Income	0.2182	0.0760	0.004	0.5059	0.1490	0.001
Constant	-9.1036	1.6233	0.000	-13.0496	3.4675	0.000
# observations	216076					
χ^2_{50}	1273.18		0.000			

*CF = control function.

Table 5: Probability of staying poor (transitions from IP)

From Insufficient Protection to	Insufficient Protection			Social Assistance		
Variable	b	$s.e.$	p	b	$s.e.$	p
Working	-0.7817	0.1705	0.000	-1.8124	0.7628	0.018
Higher sec. educ.	-0.2540	0.2319	0.273	-0.2045	0.7349	0.781
Higher education	-0.9230	0.4014	0.021	-1.5419	1.3996	0.271
Self-employed	1.5282	0.1749	0.000	0.3162	1.3140	0.810
Gender	0.0948	0.1111	0.393	0.6137	0.5714	0.283
Cohabiting	-0.2114	0.1271	0.096	-1.5415	0.6967	0.027
Household size	-0.1092	0.0540	0.043	-1.7022	0.5763	0.003
# kids $<$ 12 y.	0.0682	0.0819	0.405	2.0163	0.8038	0.012
# kids 12-16 y.	0.5725	0.1804	0.002	2.5879	0.5473	0.000
Bad health	0.0083	0.0659	0.900	0.4244	0.2511	0.091
non-EU citizen	0.2135	0.2507	0.394	-1.2078	1.4321	0.399
City	-0.2031	0.1230	0.099	0.3718	0.5684	0.513
Brussels region	0.3268	0.1998	0.102	-0.7844	0.9231	0.395
Walloon region	0.1488	0.1187	0.210	-0.1937	0.4833	0.689
Ec. growth	-0.1463	0.0531	0.006	-0.3802	0.4616	0.410
Unempl. rate	0.0187	0.2060	0.928	1.2859	1.8486	0.487
CF schooling	0.2062	0.1114	0.064	-0.2044	0.2718	0.452
CF init. cond work	-0.0192	0.0471	0.683	0.0171	0.1971	0.931
CF work	-0.4631	0.0906	0.000	-0.0258	0.2096	0.902
CF init. cond. Insuff. prot.	0.0550	0.0329	0.094	0.0404	0.1525	0.791
CF init. cond. Min. Income	0.0753	0.0685	0.272	-0.0146	0.2209	0.947
Constant	2.6787	1.9033	0.159	-12.3578	17.0562	0.469
# observations	4596					
χ^2_{42}	712.59		0.000			

*CF = control function.

Table 6: Probability of staying poor (transitions from MI to IP or MI)

From Social Assistance to	Insufficient Protection			Social Assistance		
Variable	b	$s.e.$	p	b	$s.e.$	p
Working	1.0342	1.3064	0.429	-0.8780	0.3679	0.017
Higher sec. educ.	0.3080	1.6194	0.849	-0.4266	0.4362	0.328
Higher education	-2.0154	2.2135	0.363	-1.3658	0.7674	0.075
Gender	-0.1556	0.8118	0.848	-0.3913	0.2771	0.158
Cohabiting	-0.9151	1.5123	0.545	-0.9942	0.2515	0.000
Household size	-1.0780	0.9879	0.275	0.3051	0.1607	0.058
# kids < 12 y.	0.7464	1.0758	0.488	-0.2281	0.2198	0.299
# kids 12-16 y.	1.4485	1.0238	0.157	0.7064	0.3480	0.042
Bad health	0.1887	0.4186	0.652	0.1842	0.1425	0.196
non-EU citizen	0.5938	2.1755	0.785	0.7429	0.5487	0.176
City	0.4543	0.8566	0.596	-0.2625	0.2929	0.370
Brussels region	-0.2361	1.7224	0.891	0.1962	0.4855	0.686
Walloon region	-0.9390	0.9126	0.303	0.1111	0.3030	0.714
Ec. growth	0.2841	0.3467	0.413	0.1432	0.1226	0.243
Unempl. rate	-1.7234	1.2064	0.153	-0.9188	0.4832	0.057
CF schooling	0.1015	0.5449	0.852	0.1076	0.1887	0.568
CF init. cond work	-0.7033	0.4855	0.147	-0.0946	0.1243	0.447
CF work	-0.8030	0.2151	0.000	-0.5761	0.1133	0.000
CF init. cond. Insuff. prot.	0.7683	0.4549	0.091	0.3826	0.2206	0.083
CF init. cond. Min. Income	-0.3439	0.3288	0.296	0.1966	0.0847	0.020
Constant	15.0182	11.1717	0.179	10.4613	4.4097	0.018
# observations	1703					
χ^2_{40}	558.63		0.000			

*CF = control function.

poor. It also lowers the latter risk by increasing the probability of getting and keeping a job.

- Women obtain higher degrees, but tend to lose this advantage through lower employment probabilities.
- Younger people enter the labour market better educated, but experience more fluctuations. Even after controlling for employment, low age entails a higher risk of getting poor, possibly due to lower wages and social security benefits. This also suggests that better education does not protect younger people completely from poverty.
- Family composition: singles not only find it harder to get a job, but are also more vulnerable in other ways. For example, they face relatively higher fixed expenses relative to their income. The presence of younger children not only lowers employment possibilities, but also constitutes a heavy burden on the family budget, thus generating a twofold poverty risk.
- Being a foreigner (especially from outside the EU) lowers the probability of finding and keeping a job. It also lowers family income (after controlling for employment) and the probability of receiving social assistance.
- Poor health mainly affects employment probabilities, but does not seem to have much direct effect on the likelihood of falling onto poverty.
- City dwellers are characterised by a higher labour market volatility and higher social assistance uptake. There also seem to be regional differences in terms of protection through the minimum income. Flemish people run a higher risk of IP and have a lower probability of benefiting from social assistance. Whether this pattern results from higher informal solidarity or voluntary non-take up in Flanders or from greater generosity of social services in Brussels and Wallonia, is unclear.
- The effects of the economy at large are at least dubious. A higher country-wide unemployment rate increases inflow rates into and decreases exit probability from IP, but does not seem to have any effect on transition probabilities to and from MI.

3 Microsimulation of Policies

Every anti-poverty policy presumably has a different impact on the transition probabilities between the three poverty states. In the this section, we will examine the effects of three broad categories of policies, by means of ex-post micro-simulation of some typical examples of measures:

1. optimization of the coverage of social assistance: every household which becomes poor will get social assistance,
2. activation: a temporary job is offered to all jobless poor individuals,
3. education: low-skilled individuals are encouraged to obtain a diploma of upper secondary education.

Each of these strategies can be seen as representing one of three competing views on the welfare state: the traditional welfare state, the active welfare state or the knowledge-based society.

In our simulations we assume that the effects of each strategy apply as from January 1993, the beginning of our observation period. We will indeed apply ex-post microsimulation (Merz (1991)): each policy will be applied to each member of their respective target groups present in our database. This procedure allows us to compare the different policies without having to generate hypothetical macro-economic time series nor representative sample individuals.

The target groups consist of people to whom the conditions of the specific policy apply in January 1993. We will simulate the policies for these groups only and we do not consider 'late joiners' into the respective programs. For each individual, we know the starting poverty and employment states, or we can predict them using the static estimations for the initial period. We also know, for each individual, the labour market and poverty transition probabilities[6], which allows us to construct a time path of probabilities for both employment and poverty states. Comparing time-paths with and without policy intervention gives an indication of the impact of this policy over time.

3.1 Full Coverage of the Guaranteed Minimum Income

Under this scenario, everybody in IP in January 1993 receives social assistance. We assume that reception of income support entails behavioral changes: conditional on observed characteristics, our target group will adopt the transition patterns of the MI group. The target group in our sample consists of 170 people in the IP state in January 1993.

In Figure 1 three time-paths of the probability of being non-poor are plotted: the crosses depict the observed probabilities in the target group, the grey circles represent the simulated time-path without any program and the diamonds describe the simulated time-path after application of the program. Visual comparison of the observed with the simulated baseline informs us that the predictions (and thus the estimations) are quite accurate. We also notice that spontaneous exit out of poverty amounts to 80% after approximately 2 year.

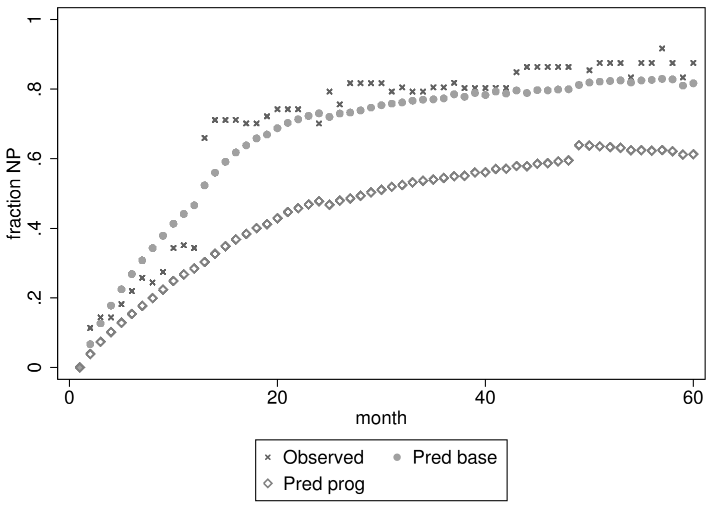
A successful policy should achieve the following goals:

[6] When a policy affects the labour market transitions, the poverty transition probabilities are obtained by predicting them with and without employment and then mixing them with the employment probabilities.

1. accelerate the spontaneous exit,
2. lift the remaining 20% out of poverty,
3. prevent new entries into poverty.

Increased coverage achieves none of the three goals above: it decelerates the spontaneous exit and decreases it to about 60% over five years. This approach is purely curative: it does not alter the entry probability, but alleviates its effect. The potentially positive effects on the exit probabilities of social assistance are clearly offset by a 'poverty trap' effect.

Figure 1: Full coverage of the guaranteed minimum income: predicted effects on the time path of (non-)poverty



The impact of increased coverage can also be illustrated in a different way. Knowledge of the transition probabilities allows us to compute some steady-state parameters for each individual, which, averaged over the target group, are given in Table 7. For comparison, these numbers are also given for the total sample. The probability of being poor for the population out of school and not yet retired amounts to 3.15%, two thirds of which do not apply for social assistance. The mean duration of a spell in poverty is about 8 months. Looking at the target group, the picture changes drastically. In the long run and without extra policy measure, about 21% of the target group live on or below the poverty threshold, with a mean spell of 13 months in IP and slightly more than 3 years in MI. A policy of 100% coverage by social assistance would *raise* the probability of poverty to 36.6% and the expected duration to almost 6.5 years.

These findings do sound somewhat paradoxical: strengthening the safety net raises the poverty risk. Of course this conclusion follows directly from the yardstick with which we

Table 7: Steady state characteristics of the full coverage scenario

	Total Sample	Target Group	
		no program	100% coverage
$\hat{\Pr}$ [NP]	96.85	79.34	63.44
$\hat{\Pr}$ [IP]	2.22	15.75	0.0
$\hat{\Pr}$ [SA]	0.93	4.91	36.56
$\hat{E}[t_{NP}]$	1117	116.9	116.9
$\hat{E}[t_{IP}]$	7.43	13.31	0.0
$\hat{E}[t_{SA}]$	8.42	37.35	77.81

chose to measure the effects of a policy. In no way do we advocate the abolishment of social assistance, which at least fills income gaps and thereby reduces the severity of poverty. On the other hand, this exercise also shows the potentially adverse effects on the poverty *dynamics* of an increased social assistance coverage.

3.2 Activation

In this scenario, the unemployed poor get a job for a period of one year. A first expected, direct effect is that this job will increase the exit probability from, and lower the (re-)entry hazard into poverty. A second, indirect effect is that persistence in employment will sustain this effect after the end of the program.

The target group in our sample consists of 160 individuals in January 1993, who are offered and supposed to accept a job at that moment. Without any program, about 30% of the poor unemployed manage to be at work after 5 years. A first direct effect of the activation policy is that the estimated probability of being at work rises by about 5.5% four years after the program is finished. Of the participants, however, more than 60% become unemployed again.

In reality, the poverty alleviation effects of this policy will strongly depend on the quality and the suitability of the job offered, parameters unaccounted for in this simulation. By setting the 'at work'-dummy equal to one, we implicitly assume that the program provides jobs of the same quality as those that are otherwise performed voluntarily by persons with comparable characteristics (except for the duration which is kept fixed here).

The activation policy (see Figure 2) now seems to affect mainly the *timing* of poverty exits. The direct effects are (a) a substantial increase of the exit and (b) a decrease of the entry probability. After 12 months the program reaches its maximum impact: it lifts an extra 23% of the participants above the poverty line, compared to the trend without intervention. Later on this result diminishes as the policy reaches its ceiling while the baseline poverty odds keep diminishing. The net residual effect of this program is about 3.7% four years after its termination. The modest long-term residual effect of this policy can also be noticed from

Figure 2: Activation scenario: anti-poverty effects

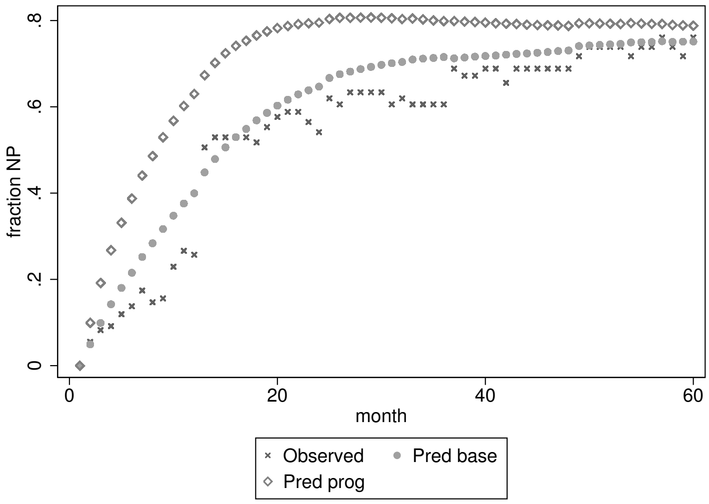


Table 8: Steady state characteristics of the activation scenario

	no program	activation
$\hat{\Pr}$ [NP]	73.47	77.90
$\hat{\Pr}$ [IP]	14.25	11.38
$\hat{\Pr}$ [SA]	12.28	10.73
$\hat{\mathrm{E}}\,[t_{NP}]$	107.4	134.8
$\hat{\mathrm{E}}\,[t_{IP}]$	12.30	9.48
$\hat{\mathrm{E}}\,[t_{SA}]$	45.87	38.04

Table 8: the steady-state probability of being non-poor increases by about 3.5%. The mean spell out of poverty, however, increases from 9 to 11.5 years.

3.3 Education

The most recent welfare state paradigm stresses education and knowledge as determining factors of social integration. We translate this into a scenario where the lowest-skilled are encouraged to obtain a degree of upper secondary education. In the 'youth variant', the target group consists of all low-skilled below the age of 25, in the 'learn-fare' variety it is made up of the poor low-skilled younger than 50 years.

The first target group in our sample consists of 192 individuals who, in January 1993, are younger than 25 and have no degree of upper secondary education. Of these 179 (93.23%) are not poor, 4 (2.08%) are insufficiently protected and 9 (4.69%) receive social assistance. The

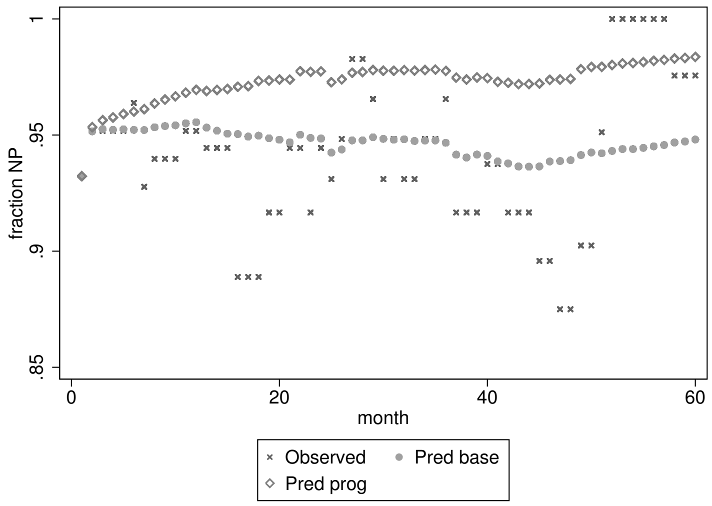


Figure 3: Education scenario, target group < 25 years: anti-poverty effects

small number of low-skilled school-leavers living below the poverty line can be explained by the fact that most of them still live with their parents. Some form of protection seems to spring from their social capital. However, since we consider a period of 5 years, our model should implicitly account for the period in which these youngsters leave their parental household to live on their own. Nevertheless, a degree of upper secondary education seems to offer some extra and lasting protection against poverty of hardly 5%.

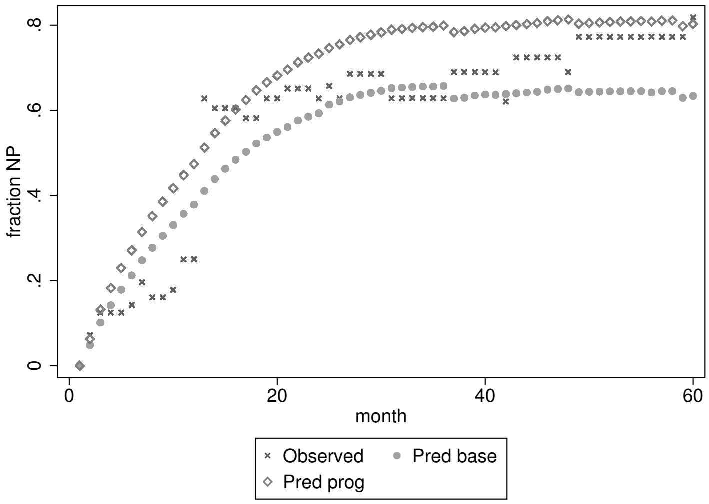
In the second variant (learnfare), the target group consists of 67 low-skilled poor respondents below the age of 50. Figure 4 again shows the lasting effects of increased education. The probability of being above the poverty line increases by 17%. This relatively large impact is also reflected in a 40% decrease of the mean spell duration in social assistance.

3.4 Policy implications

Despite the methodological and data problems discussed in sections 1 and 2, the following conclusions seem to emerge from our analysis.

- Raising the *coverage of social assistance*, while alleviating the harshest effects of poverty, also tends to increase the number of poor through the poverty trap effect. Admittedly, the findings relate to the period 1993-1997 in Belgium, in the context of a sluggish economy and rather 'passive' income compensation policies. In the mean time, work incentives have been built into the social assistance regulations and benefits have been linked with activation. Nevertheless, the simulation warns against the possible perverse effects of mere income compensation.

Figure 4: Education scenario, target group poor < 50 y: anti-poverty effects



- Getting people into *work* for a limited period (one year) affects mainly the timing of the poverty exits, but has less effect on the steady state parameters. Exits from poverty accelerate in the short short run. However, the longer-term impact of activation is very modest unless high-quality jobs are offered (e.g. combinations of work and training).

- The *education* scenario appears to yield the most substantial and durable effects, especially when focused on those living in poverty (learnfare variant).

In order to make the estimated effects of the three strategies comparable, we have to take into account the size differences of the initial target groups. To do so we reweighed the reported results. In Table 9 the poverty impact of policies is reported as a percentage of the overall group of poor people (IP and MI) in the initial period (January 1993). Raising the coverage of social assistance will increase the steady state fraction of poor people by up to 3.94%. This result is the net effect of the decreased share of under-protected people (3.90%) and the increased share of people receiving social assistance (7.84%).

Upon comparing the strategies with each other, increasing the coverage of MI has (by definition) the highest impact on extreme poverty (IP), while improving education of the low-skilled has the highest overall impact (IP and MI), strikingly more than activation of the unemployed. And yet, all in all, none of the simulated strategies appear to provide the panacea against poverty.

The rather modest impact of our simulations is of course partly due to the criterion we use to compare the different policies (fraction of the initially poor who remain poor in the long run), which is of course affected by a sizeable deadweight effect. The negative connotation of

Table 9: Net steady-state impact of policies as a fraction of the total number of poor in the initial period

	NP	IP	MI
100% coverage	-3.94%	-3.90%	+7.84%
Activation	+3.14%	-2.03%	-1.10%
Education of the young	+1.22%	-0.14%	-1.08%
Education of the poor	+5.11%	-1.09%	-4.03%

the latter term seems rather unfair, since the fact that so many initially poor finally escape poverty reflects both the effectiveness of other existing poverty-alleviating measures and the adaptability of human nature to difficult conditions.

4 Conclusion and Further Research

In this paper we model employment and poverty states as a discrete first-order Markov process, taking into account endogeneity of schooling. Using this model, we then evaluate the impact of different policies on the poverty dynamics by ex-post microsimulation. The policies we assess are exemplars of currently prevailing social policy paradigms: the traditional welfare state, the active welfare state or the knowledge-based society. When the *duration of poverty spells* or the *probability of being poor* is taken as the criterion for evaluation, our research indicates that (a) increased coverage of the guaranteed minimum wage has adverse effects, (b) activation has large short-term and small long-term positive effects, and (c) learnfare has lasting positive effects.

At the same time, our dynamic approach proved to be much more realistic in predicting the impact of policies. Given the high degree of mobility into and out of poverty, the net effects of anti-poverty measures appear to be much smaller than a static model would predict. Moreover, depending on the type of policy adopted, long-term effects may be much greater or smaller than short-term effects.

References

[1] Alcock, P. (2004), "The Influence of Dynamic Perspectives on Poverty Analysis and Anti-Poverty Policy in the UK", *Journal of Social Policy*, **33**(3), p.395-416.

[2] Bane, M.J., Ellwood, D.T. (1986), "Slipping Into and Out of Poverty: The Dynamics of Spells", *Journal of Human Resources*, **21**(1), p1-21.

[3] Böheim, R., Ermisch, J.F., Jenkins, S.P. (1999), "The Dynamics of Lone Mother's Incomes: Public and Private Income Sources Compared", *ISER Working Paper*, **1999-05**, University of Essex, UK.

[4] Breen, R., Moisio, P. (2004), "Poverty Dynamics Corrected for Measurement Error", *Journal of Economic Inequality*, **2**(3), 171-191.

[5] Callens, M. (2004), "Poverty Dynamics in Europe. A Multilevel Discrete-Time Recurrent Hazard Analysis", Chapter 4 in *Essays on Multilevel Logistic Regression*, Unpublished Ph.D. Thesis, K.U.Leuven.

[6] Cantó, O. (2000), "Income Mobility in Spain: How Much Is There?", *Review of Income and Wealth*, **46**(1), p.85-102.

[7] Cantó, O. (2002), "Climbing Out of Poverty, Falling Back In: Low Income Stability in Spain", *Applied Economics*, **34**(15), p.1903-1916.

[8] Cappellari, L. (2001), "Earnings Mobility among Italian Low Paid Workers", *ISER Working Paper*, **2001-13**, University of Essex.

[9] Cappellari, L., Jenkins, S.P. (2002), "Who Stays Poor? Who Becomes Poor? Evidence from the British Household Panel Survey", *Economic Journal*, **112**(478), p.C60-C67.

[10] Cox, D.R., Snell, E.J. (1968), "A General Definition of Residuals", *Journal of the Royal Statistical Society B*, **30**(2), p.248-275.

[11] De Blander, R., Nicaise, I. (2005), *Maatschappelijke keuzen, structurele armoede en sociale kost*, HIVA, K.U.Leuven.

[12] Devicienti, F. (2001), "Poverty Persistence in Britain: A Multivariate Analysis using the BHPS, 1991-1997", *ISER Working Paper*, **2001-02**, University of Essex.

[13] Dewilde, C. (2003), "A Life-Course Perspective on Social Exclusion and Poverty", *British Journal of Sociology*, **54**(1), p.109-128.

[14] Dewilde, C. (2004), "Poverty Mobility in the Belgian and British Welfare Regimes: The Impact of Demographic and Labour Market Events", *mimeo*, RC28 Spring Meeting on Social Stratification, Mobility, and Exclusion, 7-9 May 2004, Neuchâtel, Switzerland.

[15] DiPrete, T.A., McManus, P.A. (2000), "Family Change, Employment Transitions, and the Welfare State: Household Income Dynamics in the United States and Germany", *American Sociological Review*, **65**(3), p.343-370.

[16] Dubin, J.A., McFadden, D.L. (1984), "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption", *Econometrica*, **52**(2), p.345-362.

[17] Finnie, R., Sweetman, A. (2003), "Poverty Dynamics: Empirical Evidence for Canada", *Canadian Journal of Economics-Revue canadienne économique*, **36**(2), p.291-325.

[18] Gong, G., Samaniego, F.J. (1981), "Pseudo Maximum Likelihood Estimation: Theory and Applications", *Annals of Statistics*, **9**(4), p.861-869.

[19] Gouriéroux, C., Monfort, A., Renault, E., Trognon, A. (1987), "Generalized Residuals", *Journal of Econometrics*, **34**(1-2), p.5-32.

[20] Gustafsson, B, Muller, R., Negri, N., Voges, W. (2002), "Paths Through (and Out of) Social Assistance", in Saraceno, C. (ed.), *Social Assistance Dynamics in Europe. National and Local Poverty Regimes*, The Policy Press, Bristol.

[21] Heckman, J.J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *The Annals of Economic and Social Measurement*, **5**(4), p.475-492.

[22] Heckman, J.J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", *Econometrica*, **46**(4), p.931-959.

[23] Heckman, J.J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, **47**(1), p.153-161.

[24] Heckman, J.J. (1981a), "Statistical Models for Discrete Panel Data", in Manski, C. and McFadden, D. (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, p.114-178.

[25] Heckman, J.J. (1981b), "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process", in Manski, C. and McFadden, D. (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, p.179-195.

[26] Jenkins, S.P. (2000), "Modelling household income dynamics", *Journal of Population Economics*, **13**(4), p.529-567.

[27] Merz, J. (1991), "Microsimulation - a Survey of Principles, Developments and Applications", *International Journal of Forecasting*, **7**, p.77-104.

[28] Mortelmans, D., Casman, M.T., Doutrelepont, R. (eds.) (2004), *Elf jaar uit het leven in België. Socio-economische analyses op het Gezinsdemografisch Panel PSBH*, Academia Press, Ghent, Belgium.

[29] Murphy, K.M., Topel, R.H. (1985), "Estimation and Inference in Two-Step Econometric Models", *Journal of Business and Economic Statistics*, **3**(4), p.370-379.

[30] Nicaise, I., Groenez, S., Adelman, L., Roberts, S., Middleton, S. (2004), *Gaps, Traps and Springboards in European Minimum Income Systems*, Leuven: HIVA / Loughborough: CRSP (2 vol.), 134+268p.

[31] Noble, M., Cheung, S.Y., Smith, G. (1998), "Origins and Destinations - Social Security Claimant Dynamics", *Journal of Social Policy*, **27**(3), p.351-369.

[32] Parke, W.R. (1986), "Pseudo Maximum Likelihood Estimation: The Asymptotic Distribution", *Annals of Statistics*, **14**(1), p.355-357.

[33] Pierce, D.A. (1982), "The Asymptotic Effect of Substituting Estimators for Parameters in Certain Types of Statistics", *Annals of Statistics*, **10**(2), p.475-478.

[34] Randles, R.H. (1982), "On the Asymptotic Normality of Statistics with Estimated Parameters", *Annals of Statistics*, **10**(2), p.462-474.

[35] Riphahn, R.T. (2001), "Rational Poverty or Poor Rationality? The Take-Up of Social Assistance Benefits", *Review of Income and Wealth*, **47**(3), p.379-398.

[36] Stevens, A.H. (1999), "Climbing Out of Poverty, Falling Back In - Measuring the Persistence of Poverty Over Multiple Spells", *Journal of Human Resources*, **34**(3), p.557-588.

[37] Stewart, M.B., Swaffield, J.K. (1999), "Low Pay Dynamics and Transition Probabilities", *Economica*, **66**, p.23-42.

[38] Vella, F., Verbeek, M. (1999), "Two-Step Estimation of Simultaneous Equation Panel Data Models with Censored Endogenous Variables", *Journal of Econometrics*, **90**(2), p.239-263.

[39] White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models.", *Econometrica*, **50**(1), p.1-25.

A Appendix

A.1 Generalized Residuals

A.1.1 Schooling Equation

Defining $\Lambda(q) = \frac{e^q}{1+e^q}$, it can be shown that

$$\mathrm{E}\left[u_i \mid x_i, s_i = 1\right] = (\mu_1 - \alpha' x_i) - \frac{\ln\left(1 + e^{\mu_1 - \alpha' x_i}\right)}{\Lambda(\mu_1 - \alpha' x_i)},$$

$$\mathrm{E}\left[u_i \mid x_i, s_i = 2\right] = \frac{\ln\left(\frac{1+e^{\mu_1 - \alpha' x_i}}{1+e^{\mu_2 - \alpha' x_i}}\right) + (\mu_2 - \alpha' x_i)\Lambda(\mu_2 - \alpha' x_i) - (\mu_1 - \alpha' x_i)\Lambda(\mu_1 - \alpha' x_i)}{\Lambda(\mu_2 - \alpha' x_i) - \Lambda(\mu_1 - \alpha' x_i)},$$

and

$$\mathrm{E}\left[u_i \mid x_i, s_i = 3\right] = \frac{\ln\left(1 + e^{\mu_2 - \alpha' x_i}\right) - (\mu_2 - \alpha' x_i)\Lambda(\mu_2 - \alpha' x_i)}{1 - \Lambda(\mu_2 - \alpha' x_i)}.$$

A.1.2 Work Equation

Defining $\tilde{u}_i = \mathrm{E}\left[u_i \mid x_i, s_i\right]$ and $\tilde{v}_{i1} = \mathrm{E}\left[v_{i1} \mid y_{i1}, s_i, w_{i1}\right]$, we have

$$\mathrm{E}\left[v_{it;w_{i;t-1}} \mid y_{it}, s_i, w_{it} = 0\right] = -\frac{\left(m_{it;w_{i;t-1}}\right)\Lambda\left(-m_{it;w_{i;t-1}}\right) + \ln\left(1 + e^{-m_{it;w_{i;t-1}}}\right)}{\Lambda\left(-m_{it;w_{i;t-1}}\right)}$$

and

$$\mathrm{E}\left[v_{it;w_{i;t-1}} \mid y_{it}, s_i, w_{it} = 1\right] = \frac{\left(m_{it;w_{i;t-1}}\right)\Lambda\left(-m_{it;w_{i;t-1}}\right) + \ln\left(1 + e^{-m_{it;w_{i;t-1}}}\right)}{1 - \Lambda\left(-m_{it;w_{i;t-1}}\right)},$$

where

$$m_{it;w_{i;t-1}} = \beta_1' y_{it} + \beta_2 d_{s_2;i} + \beta_3 d_{s_3;i} + \beta_4 \tilde{u}_i,$$

for the first period, and

$$m_{it;w_{i;t-1}} = \beta_{1;w_{i;t-1}}' y_{i1} + \beta_{2;w_{i;t-1}} d_{s_2;i} + \beta_{3;w_{i;t-1}} d_{s_3;i} + \beta_{4;w_{i;t-1}} \tilde{u}_i + \beta_{5;w_{i;t-1}} \tilde{v}_{i1},$$

for the other periods.

A.1.3 Initial Poverty Equation

$$\mathrm{E}\left[\varepsilon_{i1;j} \mid z_{i1}, s_i, w_{i1}, \tilde{u}_i, \tilde{v}_{i1}; p_{i1} = k\right] = \begin{cases} -\frac{\lambda\sqrt{3}}{\pi} \ln\left[\Lambda\left(n_{i1;k}\right)\right] & \text{if } i = j, \\ \frac{\lambda\sqrt{3}}{\pi} \frac{\Lambda(n_{i1;j})}{1-\Lambda(n_{i1;j})} \ln\left[\Lambda\left(n_{i1;j}\right)\right] & \text{if } i \neq j, \end{cases}$$

with

$$n_{i1;k} = \gamma_{1;k}' z_{i1} + \gamma_{2;k} d_{s_2;i} + \gamma_{3;k} d_{s_3;i} + \gamma_{4;k}' w_{i1} + \gamma_{5;k} u_i + \gamma_{6;k} v_{i1},$$

see Dubin and McFadden (1985).

B Correction of the Asymptotic Covariance-matrix for Generated Regressors[7]

Consider a multivariate model with log-likelihood equal to $f\left(y_i, z_i \mid x_i; \alpha, \beta\right) = g\left(y_i \mid z_i, x_i; \alpha, \beta\right) h\left(z_i \mid x_i; \beta\right)$. Maximize now first $\sum_i \ln h\left(z_i \mid x_i; \beta\right)$ with respect to β, and then $\sum_i \ln g\left(y_i \mid z_i, x_i; \alpha, \hat{\beta}\right)$ with

[7] See for instance Gong and Samaniego (1981), Murphy and Topel (1985), Parke (1986), Pierce (1982), Randles (1982) and Vella and Verbeek (1999).

respect to α. Define now the following matrices

$$H_{11} = \mathrm{E}\left[-\frac{\partial^2 \ln g}{\partial\alpha\partial\alpha'}\right]$$

$$\begin{aligned} H_{12} &= \mathrm{E}\left[-\frac{\partial^2 \ln g}{\partial\alpha\partial\beta'}\right] \\ &= \mathrm{E}\left[-\frac{\partial^2 \ln g}{\partial\alpha\partial\lambda'}\frac{\partial\lambda}{\partial\beta'}\right], \end{aligned}$$

with λ the generalized residual from the estimation of β by maximizing $\sum_i \ln h\left(z_i \mid x_i;\beta\right)$. The Taylor series of $\partial \ln g\left(\hat{\alpha},\hat{\beta}\right)/\partial\alpha$ around (α,β) is given by

$$\begin{aligned} N^{-1}\sum_{i=1}^{N}\frac{\partial \ln g\left(\hat{\alpha},\hat{\beta}\right)}{\partial\alpha} &= N^{-1}\sum_{i=1}^{N}\frac{\partial \ln g\left(\alpha,\beta\right)}{\partial\alpha} + N^{-1}\sum_{i=1}^{N}\frac{\partial^2 \ln g\left(\alpha,\beta\right)}{\partial\alpha\partial\alpha'}\left(\hat{\alpha}-\alpha\right) \\ &\quad + N^{-1}\sum_{i=1}^{N}\frac{\partial^2 \ln g\left(\alpha,\beta\right)}{\partial\alpha\partial\beta'}\left(\hat{\beta}-\beta\right) + o_p\left(1\right). \end{aligned}$$

At the optimum, $\hat{\alpha}$, the LHS is equal to zero, resulting in

$$\sqrt{N}\left(\hat{\alpha}-\alpha\right) = H_{11}^{-1}\left\{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\partial \ln g}{\partial\alpha} + H_{12}\sqrt{N}\left(\hat{\beta}-\beta\right)\right\} + o_p\left(1\right).$$

Asymptotic independence of the first two terms on the RHS implies

$$\sqrt{N}\left(\hat{\alpha}-\alpha\right) \sim N\left(0; H_{11}^{-1}\left\{H_{11} + H_{12}\Sigma_\beta H_{12}'\right\}H_{11}^{-1}\right). \tag{8}$$

Estimation of α by pseudo-maximum likelihood, compels us to adapt (8) as follows

$$\sqrt{N}\left(\hat{\alpha}-\alpha\right) \sim N\left(0; H_{11}^{-1}\left\{Q_{11} + H_{12}\Sigma_\beta H_{12}'\right\}H_{11}^{-1}\right), \tag{9}$$

where $Q_{11} = \mathrm{E}\left[\frac{\partial \ln g}{\partial\alpha}\cdot\frac{\partial \ln g}{\partial\alpha'}\right]$. The quantities H_{11}, H_{12} and Q_{11} can be consistently estimated by replacing expectations by sample means.