A Dynamic Discrete Choice Model of Labor Supply

Peter Haan

(DIW Berlin & FU Berlin)

16.03.2005

Introduction I

- There exists an important literature on estimating labor supply using a discrete choice specification (van Soest, 1995).

- At the same time there are numerous studies on the dynamics of labor supply (Blundell and MaCurdy, 1999).

- In this paper, I try to combine this literature by developing a dynamic discrete choice model.

- This framework has the advantage that on the one hand nonlinearities in the budget set can be modelled and, on the other hand, frictions of the labor market are included via a partial adjustment mechanism.

Introduction II

- The empirical analysis is based on a dynamic conditional logit model with random effects:
  - unobserved heterogeneity is modelled in nonparametric way
  - problem of initial condition is taken into account following Wooldridge (2005)

- Based on panel data from the SOEP, I estimate the dynamic discrete choice model for single men and single women in Germany for the period 1999-2002.

- Short and long term labor supply elasticities are derived using the Markov chain property.

- Main finding: in the long run, labor supply adjustment of households triples relative to the short run.

Theoretical Background I

- Discrete choice models are based on the assumption that a household (i) is faced with a finite number (J) of discrete bundles of leisure and income, which provide different levels of utility (U_j) at period t.

$$U_{ijt} = U(X_{it}, ls_{ijt}, y_{ijt}, c_i, \epsilon_{ijt}) \tag{1}$$

- Utility depends on leisure (ls_{ijt}) and net household income (y_{ijt}), which differ over the alternatives, on individual variables (X_{it}) and on an unobservable component consisting of an individual specific term (c_i) and of an error term that varies over time, individuals and alternatives ϵ_{ijt}.

Theoretical Background II

- In the static model of labor supply households do not face other constraints than the budget function, therefore the decision rule is:

$$Pr_{itk} = Pr(U(X_{it}, ls^u_{ikt}, y^u_{ikt}, c_i, \epsilon_{ikt}) > \\ U(X_{it}, ls^u_{imt}, y^u_{imt}, c_i, \epsilon_{imt})); \quad \forall m \neq k \tag{2}$$

- Due to changes in a household‘s budget function or due to changes of characteristics it might become optimal for the household to adjust labor supply over time.

- In this model a household does not face any frictions and can therefore adjust labor supply immediately: hence, the household is at any point time in his optimal alternative.

Theoretical Background III

- In contrast to the static model, in the dynamic framework, frictions of the labor market are included. Frictions are modelled via a partial adjustment mechanism:

$$h_{it} - h_{it-1} = \theta(h_{it}^{u} - h_{it-1}); 0 \leq \theta \leq 1 \tag{3}$$

- Therefore, in this model unconstrained labor supply in period t has the following form:

$$h_{it}^{u} = f(h_{it}, h_{it-1})$$

- Hence, in the the dynamic model the decision rule is:

$$Pr_{itk} = Pr(U(X_{it}, ls_{ikt}, ls_{it-1}, y_{ikt}, c_i, \epsilon_{ikt}) > \tag{4}$$
$$U(X_{it}, ls_{imt}, ls_{it-1}, y_{imt}, c_i, \epsilon_{imt})); \quad \forall m \neq k$$

Data Organization

- Population: Single households, aged 20-65, not self employed, not in education, not retired.

- Data: Socio Economic Panel (2000-2003)

- The dependent variable y_t is an indicator assigning households to their chosen working hours category j in period t.

$$y_t = \begin{cases} 1 & \text{if } y_t \in j, \text{where j = 1,2,...,J} \\ 0 & \text{otherwise} \end{cases}$$

- The key explanatory variable, the disposable net household income is derived on basis of the microsimulation model STSM.

	women			men		
alternatives	hours	%	income	hours	%	income
no-work	0	18.67	832	0	15.68	644
part-time1	15	12.14	1212	-	-	-
part-time2	29	13.18	1415	-	-	-
full-time	38	40.01	1637	37	56.1	1689
over-time	46	15.99	1838	47	28.23	2101
age (year 2000)	43.05			42.24		
East-Germany	20.99%			25.90%		
households with childern						
younger 6	5.02%			0.12%		
previous working history						
years of full-time	14.70			17.70		
years of part-time	3.34			0.62		
observations	823			618		
period	3.6			2.7		

Econometric Specification I

- If the error terms ϵ_{ijt} follow a Gumble distribution the probability of choosing alternative k conditional on the explanatory variables, the chosen alternative of the previous period and the unobserved individual effect has the following form:

$$Pr(Y_{it} = k | X_{it}, z_{itj}, ls_{it-1}, c_i^m) = \frac{\exp U(X_{it}, ls_{ikt}, ls_{it-1}, y_{ikt}, c_i^m)}{\sum_{j=1}^{J} \exp U(X_{it}, ls_{ijt}, ls_{it-1}, y_{ijt}, c_i^m)}$$

- Unobserved heterogeneity is specified in a nonparametric way. The unobserved heterogeneity is described by an arbitrary discrete probability distribution $P_i(c_i^m)$ with a small number of mass points $c_i^m, \forall m (m = 1, 2, ...M)$.

Econometric Specification II

- The log-likelihood to be maximized is then:

$$l = \sum_{i=1}^{n} \sum_{m=1}^{M} P_i(c_i^m) \sum_{t=1}^{T_i} \sum_{j=1}^{J} d_{itj} Pr(y_{it} = j) \tag{5}$$

- Estimation of the likelihood function would produce inconsistent estimates as the lagged dependent variable is correlated with the individual specific effect (Initial condition problem) .

Econometric Specification III

- In order to solve this problem, I follow Wooldridge (2005) and assume that there exists a correctly specified model for the unobserved individual effect $h(c|y_0, x; \delta)$ conditional on the initial state (y_0) and on individual specific variables that are constant over time (x) .

- Further, I assume a normal distribution of $h(c|y_0, x; \delta)$, which leads to a linear specification of c_i:

$$c_i = \alpha_0 + \alpha_1 y_{i0} + x_i \alpha_2 + a_i \tag{6}$$

- The individual specific variables include amongst others previous working history.

Econometric Specification IV

- The utility of alternative j becomes:

$$\begin{aligned} U_{ijt} &= U(X_{it}, ls_{ijt^r}, ls_{ijt_i}, y_{ijt}, c_i(y_{i0}, x_i, a_i)), \epsilon_{ijt}) \quad (7) \\ &= U(X_{it}, ls_{ijt^r}, ls_{ijt_i}, y_{ijt}, y_{i0}, x_i, a_i, \epsilon_{ijt}) \end{aligned}$$

- That leads to a likelihood function that provides consistent estimates:

$$l = \sum_{i=1}^{n} \sum_{m=1}^{M} P(a_i^m) \sum_{t=1}^{T_i} \sum_{j=1}^{J} d_{itj} Pr(y_{it} = j) \quad (8)$$

- Maximization is based on numerical Gaussian quadrature. procedure

Econometric Specification V

- The utility function is assumed to be translog. Income and leisure, their interaction and their quadratic terms enter the utility function in logarithm.
- The individual specific variables are interacted with the logarithm of the leisure time.
- The lagged labor supply enters with a vector of dummy variables, where non participation is the base category.
- Unobserved heterogeneity is specified as a random coefficient of leisure time.

	single men		*single women*	
	coefficient	st.error	coefficient	st.error
y	8.152	3.298*	4.295	1.919*
ls	111.572	9.609**	106.510	6.796**
alt1(t-1)*ls	-	-	-1.535	0.496*
alt2(t-1)*ls	-	-	-2.070	0.453**
alt3(t-1)*ls	-2.103	0.498**	-2.102	0.348**
alt4(t-1)*ls	-3.913	0.723**	-2.429	0.516**
ls*y01	-	-	-9.439	0.950**
ls*y02	-	-	-13.836	0.973**
ls*y03	-3.077	0.549**	-17.572	1.001**
ls*y04	-8.112	1.064**	-19.553	0.998**
Mass points				
c1	-12.483	3.093**	-7.611	1.177**
c2	9.833	0.853**	1.009	.1985**
c3	0.295		14.352	
logp(c1)	-2.502	0.492**	1.141	.563*
p(c1)	0.071		0.075	
logp(c2)	-2.603	0.260**	3.633	.449**
p(c2)	0.065		0.902	
p(c3)	0.865		0.024	
Observations	1665		2989	
log likelihood	-1193.154		-3442.3303	

Note: Results of quadratic and interaction terms of income and leisure, individual specific variables and time dummies are not reported

Results II: Labor Supply Elasticities

- Labor supply elasticities are derived numerically simulating a scenario with a one 1% increase in gross hourly wages.
- The calculation of short and long term elasticities is based on the Markov chain property. For several periods transition matrices for the status quo and the simulated scenario are derived, which lead labor supply elasticities.
- Elasticities are mean elasticities of households conditional on observed and unobserved heterogeneity and the chosen labor supply alternative in period $t-1$.
- Long term elasticity are derived if households have reached their equilibrium, e.g. if the change in labor supply does not change over time.

Results III: Transition matrix for single men

	$B(0)$	$B(3)$	$B(4)$	$B(0)$	$B(3)$	$B(4)$
	all households					
	Status quo			*1% gross wage*		
$B(0)$	73.50	24.08	2.42	73.41	24.17	2.42
$B(3)$	9.35	70.86	19.79	9.31	70.85	19.84
$B(4)$	2.22	43.61	54.16	2.21	43.58	54.20

Results IV: Labor Supply Elasticities

	men (1)	men (2)	women (1)	women (2)
	change in participation (in percent)			
t	0.047	0.085	0.006	0.006
t+1	0.085	0.157	0.016	0.014
t+2	0.113	0.212	0.025	0.022
t+3	0.134	0.252	0.032	0.027
t+4	0.149	0.280	0.038	0.031
	change in hours (in percent)			
t	0.054	0.101	0.031	0.031
t+1	0.095	0.179	0.053	0.052
t+2	0.125	0.235	0.069	0.066
t+3	0.146	0.277	0.080	0.077
t+4	0.162	0.306	0.088	0.084

Note: (1) Controlling for unobserved heterogeneity (2) Not controlling for unobserved heterogeneity

Conclusion

- In this paper, I have developed a model of labor supply that includes nonlinearities in the budget set and frictions of the labor market.

- The key empirical finding is, that there exists a large difference between short and long term behavioral adjustment. The adjustment of labor supply triples in the long run.

- Further empirical results:
  - It is important to control for unobserved heterogeneity.
  - There exists a significant state dependence effect.
  - Differences of the labor supply elasticities by region and gender as well as the size of the long run elasticities is similar to the the elasticities, which have been derived within the static model.

- This study is not only interesting from a methodological point of view. Employing the dynamic model it is possible to asses the short and long term labor supply effects of policy reforms, such as a reform of income taxation.