Much Ado About Nothing: Conditional Logit vs. Random Coefficient Models for Estimating Labour Supply Elasticities *

(to appear in Applied Economics Letters)

Peter Haan

(DIW Berlin and FU Berlin) †

December 14, 2005

Abstract

This study compares several specifications of discrete choice labour supply estimations on basis of the German Socio Economic Panel. My results suggest that despite the restrictive assumptions of the error terms the conditional logit model provides an adequate model choice for the analysis of labour supply functions. Significance tests, which are based on bootstrapped confidence intervals, show that labour supply elasticities derived within the conditional logit model do not significantly differ from elasticities derived in flexible random coefficient models.

Keywords: labour supply, discrete choice models, specification test

JEL: C25, C52, J22

*Financial support by the German Science Foundation (DFG) in the priority programe "Potentials for more flexibility on heterogenous labor markets" (project STE 681/5-1) is gratefully acknowledged. I would like to thank Katharina Wrohlich, Viktor Steiner, Arne Uhlendorff, Martin Kroh, Dirk Hofmann, and seminar participants at the Nachwuchsworkshop of the German Statistical Association. The usual disclaimer applies.

†Correspondence to: Peter Haan, DIW Berlin, Königin-Luise-Straße 5, 14195 Berlin, e-mail: phaan@diw.de

1 Introduction

Estimating labour supply functions using a discrete rather than a continuous specification has become increasingly popular in recent years, as for example in van Soest (1995) and Kang et al. (2004). The main advantage of the discrete choice approach compared to continuous specifications derives from the possibility to model nonlinearities in budget functions. However, the standard discrete choice approach, the conditional logit model, is based on the restrictive assumption of homogenous error variances. This leads amongst others to the often discussed property of the independence of irrelevant alternatives (IIA) (McFadden, 1973). Econometric literature has suggested more general discrete choice models that relax the assumption of homogenous error variances and that allow for effect heterogeneity, for example the random coefficient model (Revelt and Train, 1996). However, these less restrictive specifications have shown to incur very high computational cost and to result in serious problems with maximisation. Malachow-Moeller and Svarer (2003) show that random coefficient models work well with small data sets and few alternatives. However, violating one of these conditions can lead to a serious dimensionality problem, which impedes estimation.

It is therefore of particular interest for applied research, which approach is more adequate when analysing discrete choice models: the standard conditional logit model or more general random effect models. To the extent that effect heterogeneity is present in empirical models of labour supply functions, the application of a random effect model is necessary to derive consistent estimates. However, if such heterogeneity is nonexistent or the effect heterogeneity does not have a significant impact on labour supply elasticities, standard discrete choice models provide the more favourable choice. In this paper, I will provide an empirical analysis of the two estimation procedures and will test for differences in the results. Estimations are based on the German Socio Economic Panel (SOEP).

Estimation results indicate the significant existence of unobserved heterogeneity in labour supply estimations, suggesting the application of random effect models. However, the opposite conclusion has to be drawn, when turning to labour supply elasticities, which describe the quantities implications of labour supply models most accurately. Significance tests based on bootstrapped confidence intervals reject the hypothesis that labour supply elasticities derived from conditional logit models differ from elasticities calculated in a random specification. This result is robust regardless whether the random effects are estimated parametrically or in a non parametric setting. Therefore, for computational reasons, standard discrete choice models that are more restrictive in their assumptions regarding error variances, seem to represent the adequate model choice for the analysis of labour supply models.

2 Econometric Model

Discrete choice models are based on the assumption of utility maximising behaviour of individuals. An individual i chooses among J alternatives that provide different levels of utility. The utility function U_{ij} consists of an observable part V_{ij} and random elements ϵ_{ij}:

$$U_{ij} = V_{ij} + \epsilon_{ij} \tag{1}$$

The probability that individual i chooses alternative k is:

$$Pr_{ik} = Pr(U_{ik} > U_{im}); \quad \forall m \neq k \tag{2}$$

In order to derive an operational model the crucial question is how to treat the unknown part of the utility function. McFadden (1973) showed that if (and only if) the error terms ϵ_{ij} are independently and identically distributed (iid) with type I extreme value distribution $F(\epsilon_{ij}) = \exp(-\exp(\epsilon_{ij}))$, with fixed variance $\frac{\pi^2}{6}$, the logit choice probability can be derived. Following, the probability of choosing alternative k becomes:

$$Pr_{ik} = \frac{\exp(V_{ik})}{\sum_{j=1}^{J}\exp(V_{ij})}; \quad k \in J \tag{3}$$

If the observed part of the utility function is specified to be linear in parameters, $V_{ij} = X'_{ij}\beta$, where vector X_{ij} captures K observable variables of individual i in alternative j and vector β is a vector of coefficients, the standard conditional logit model emerges. The log likelihood function to be estimated has the following form:

$$l = \sum_{i=1}^{n}\sum_{j=1}^{J} d_{ij} \ln Pr(y_i = j) \tag{4}$$

where $d_{ij} = 1$ if individual i chooses alternative j and 0 otherwise. In econometric literature, conditional logit models are often employed and their desirable properties have been widely discussed (Greene, 2003). However, the conditional logit model has severe drawbacks. Train (2003) names three main limitations of conditional logit, those being repeated choices over time, taste variation and substitution patterns. The most prominent limitation of conditional logit models resulting from the iid assumption of the error terms is the property called independence of irrelevant alternatives (IIA). This restriction implies that the odds ratio of two alternatives, j and k, does not depend on other alternatives. Hence, if the assumption of the error term distribution does not hold, the conditional logit model leads to inconsistent estimates.

In recent years several more general discrete choice models have been developed that relax the strong error term assumption and circumvent the limitations of conditional logit. Examples are generalised extreme value models, probit discrete choice models and the random coefficient model (Train, 2003). In this application, I focus on the random coefficient model, as this model is often applied, and implemented in standard software packages such as SAS, GAUSS or Stata.[1] The difference between the conditional logit model and the random coefficient model is captured in the vector of coefficients to be estimated. In the random coefficient model the coefficient vector is denoted as β_i and can be decomposed into a fixed part β and a random part μ_i:

[1]Malachow-Moeller and Svarer (2003) provide a program code for multinomial logit models with random coefficients in SAS. Train has written a program for mixed models in GAUSS. GLLAMM, developed by Rabe-Hesketh et al. (2001) allows to estimate random coefficient models in Stata. All estimations in this application have been performed using GLLAMM. I would like to thank Sophia Rabe-Hesketh for her support using this program.

$$\beta_i = \beta + \mu_i; \tag{5}$$

The random part μ_i captures non observable individual effects, such as taste, which can be modelled in a parametric or non parametric way. The researcher can neither observe nor estimate β_i. Instead the distribution of β_i has to be estimated. Theoretically, it is possible to model the random coefficient specification in a very general way by assuming all coefficients to vary randomly. However, depending on the number of coefficients this becomes enormously complex as multiple integrals have to be solved. (Train, 2003). Therefore, in this application, I assume only one of the coefficients to be random.

In the parametric case it is assumed that β_i follows some continuous distribution $f(\beta_i|\beta,\sigma)$. In most applications, β_i is assumed to be normally distributed (Train, 2003). Therefore, in the parametric random coefficient specification, the probability to choose alternative k is the integral over all possible values of β:

$$Pr_{ik} = \int_{-\infty}^{\infty} \frac{\exp(X'_{ik}\beta_i)}{\sum_{j=1}^{J}\exp(X'_{ij}\beta_i)} f(\beta_i)d(\beta_i); \quad k \in J \tag{6}$$

Heckman and Singer (1984) have derived a more flexible specification of the random coefficient model. They suggest a nonparametric method, which does not rely on a restrictive distribution assumption of β_i. Instead, it is assumed that the unobserved heterogeneity is described by an arbitrary discrete probability distribution $P_i(c^m)$ with a small number of mass points $c^m, \forall m(m = 1, 2, ...M)$, where $E(c) = \sum_{i=1}^{N}\sum_{m=1}^{M} P_i(c^m)c^m = 0$ and $\sum_{m=1}^{M} P_i(c^m) = 1$.

Mass points and their probabilities are jointly estimated with the parameters of the model using maximum likelihood. The estimation is based on the assumption that unobserved heterogeneity is independent of the explanatory variables. Note, due to the specification of the unobserved heterogeneity, only $m-1$ mass points and $m-1$ probabilities can be freely estimated. One mass point and its probability is derived according the above specified assumptions (Steiner, 2001). In a nonparametric specification, the decision rule for an individual i to choose alternative k becomes

$$Pr_{ik} = \sum_{m=1}^{M} P_i(c^m) \frac{\exp(X'_{ik}\beta_i)}{\sum_{j=1}^{J}\exp(X'_{ij}\beta_i)}; \quad k \in J \tag{7}$$

Inserting equation (6) and (7) into equation (4), the log likelihood function for the parametric and nonparametric random coefficient models can be derived. The appealing flexibility of the random specifications, which circumvent the restrictions of the conditional logit models, has enormously high computational costs. Convergence and robustness of the estimation is often problematic even if only one coefficient is specified as being random. In order to maximise the likelihood function of a random coefficient model, simulation procedures or numerical integration need to be applied.[2] Hence, relative to the conditional logit model, estimations using random specification are cumbersome. Obviously, for applied research this might be a considerable disadvantage.

[2]In this application, I employ numerical integration by Gauss-Hermite quadrature for the nonparametric model. The parametric model is estimated using adaptive Gauss-Hermite quadrature, which reduces the computational cost significantly (Rabe-Hesketh et al., 2001).

3 Empirical Analysis of Labour Supply

For the specification of the labour supply model, I follow van Soest (1995) and assume a household utility function where spouses jointly maximise utility. Household income and leisure terms of the spouses, their interaction and their quadratic terms enter the utility function in logarithm. Individual and household specific variables are interacted with the logarithm of the leisure terms and the household income.[3] Drawing on previous studies on household labour supply (Steiner and Wrohlich, 2004) I specify 13 discrete alternatives of labour supply, among which households have the choice.[4] The disposable net household income is derived on basis of the microsimulation model STSM that contains the main features of the German tax and transfer system (Haan et al., 2005). Household preferences for income and leisure might differ by individual or household specific characteristics such as age, region or health status. In this application, I focus only on married couples with both spouses having a flexible labour supply. That implies all couples are excluded in which either spouse is a civil servant, self-employed, student, on maternity leave, or retired. Only persons between 20 and 65 years of age are considered. After dropping observations due to missing variables, 2812 households remain in the sample. The year of analysis is 2001.

Inserting the household utility function into equations (3), (6) and (7), and deriving the respective likelihood functions, the conditional logit model and the

[3]For a more detailed discussion see Haan (2004).

[4]Because of the small number of men in part-time employment in the sample, only three categories could be specified for them, namely no work, full time, and overtime. For women, two additional part-time alternatives have been defined. Alternatives with too little observations are excluded. See Haan (2004) for more information about the definition of the alternatives.

parametric and nonparametric random coefficient models can be estimated. In this application, I assume the random coefficient to vary with the household income.[5]

[Table 1: about here]

Table 1 yields the results of the estimations. In the first two columns the results of the conditional logit are presented. In order to test the theoretical implications of the labour supply function, the first and second derivatives of the utility function with respect to income and leisure need to be derived (van Soest, 1995). The empirical utility function is in line with theory as all derivatives have the expected signs.

The quantitative implications of the labour supply model can best be described by deriving hours and participation elasticities with respect to given percentage change in the gross wage rate. Although a closed-form expression of elasticities is not available for the utility function estimated here, elasticities can be calculated from the simulated change in estimated hours and participation rates to an exogenous change in the gross wage rate Haan (2004). A discussion of other variables is omitted here, as this is not the focus of this research, see e.g. Steiner and Wrohlich (2004) for a detailed interpretation of the model. Before turning to the labour supply elasticities, the estimation results of the random coefficient model need to be interpreted. For the numerical integration in the parametric random coefficient model (equation 6), 10 (adaptive)

[5] Gerfin and Leu (2003) employ the same specification, van Soest (1995) allows the random effect to vary with leisure terms, whereas Duncan and MacCrae (1999) employ a random specification that varies with income and both leisure terms.

quadrature points were used. In the nonparametric specification the Akaike Criterion (AIC)[6] indicates that two mass points are required to model the unobserved heterogeneity, which has also be found by Bargain (2005). Comparing the conditional logit model to the random specifications the estimation results strongly indicate the existence of unobserved heterogeneity in the model. In both specifications the AIC suggests that the random model is superior to the conditional logit model. Further, the significant impact of the standard error in the parametric specification and the significance of the mass point in the nonparametric model support this finding. That implies that the variances of the error terms are not constant, and thus the iid assumption of the error terms is violated. Thus, the conditional logit model leads to inconsistent estimation of the coefficients.

However, this criterion is not sufficient to reject the implications of the conditional logit model. As mentioned above, labour supply elasticities provide the most adequate interpretation of discrete choice labour supply models. Therefore, I will test in the following whether the elasticities derived within the random specifications and the conditional logit model differ significantly. The test is based on bootstrapped confidence intervals of the conditional logit labour supply elasticities. The test procedure is straight forward though powerful: If the elasticities derived within the random specification fall into the 95% confidence interval of the conditional logit elasticities, the hypothesis that the elasticities do not differ significantly, can not be rejected. The following Table

[6]I follow Steiner (2001) and use the Akaike Information Criterion rather the standard likelihood ratio test, as the latter violates standard regularity conditions and its parameter distribution is not known.

2 yields average labour supply elasticities with respect to participation and with respect to working hours derived in the three models. In addition, the bootstrapped confidence intervals of the conditional logit model are presented.

[**Table 2: about here**]

The elasticities are in line with those found in previous literature, e.g. Haan and Steiner (2004). Therefore, I omit a discussion of the elasticities. The key result for my research question is that regardless of the region and of gender, all elasticities derived in both random specifications fall within the bootstrapped confidence intervals. Hence, the qualitative implication of the labour supply model resulting from the random specifications do not differ significantly from those derived within the conditional logit model.

4 Conclusion

The empirical analysis of discrete choice labour supply specification with and without random effects, has shown that despite the significant effect of unobserved heterogeneity, the implications of the labour supply models do not differ significantly between the models. That leads to the conclusion that for computational reasons the standard discrete choice model, attractive for its simple structure, provides an adequate model choice for the analysis of labour supply functions.

References

Bargain, O. (2005). On modelling labor supply with taxation. *IZA Discussion-Paper 1455.*

Duncan, A. and J. MacCrae (1999). Household labour supply, childcare cost and in-work benefits: Modelling the impact of the working families tax credit in the uk. *Manuscript, Department of Economics and Related Studies, University of York.*

Gerfin, M. and R. Leu (2003). The impact of in-work benefits on poverty and household labour supply - a simulation study for switzerland. *IZA Discussion-Paper 762.*

Greene, W. H. (2003). *Econometric Analysis* (5th ed.). Prentice Hall.

Haan, P. (2004). Discrete choice labor supply: Conditional logit vs. random coefficient models. *DIW Discussion Paper 394.*

Haan, P. and V. Steiner (2004). Distributional and fiscal effects of the german tax reform 2000 a behavioral microsimulation analysis. *DIW Discussion Paper 418.*

Haan, P., V. Steiner, and K. Wrohlich (2005). Dokumentation des steuer-transfer-mikrosimulationsmodells 1999-2002. *Mimeo.*

Heckman, J. and B. Singer (1984). A method for minimizing the distributional assumptions in econometric models for duration data. *Econometrica 52*, 271–320.

Kang, G., S. Huffman, and H. H. Jensen (2004). An empirical analysis of joint decisions on labour supply and welfare participation. *Applied Economics Letters 10*, 869—872.

Malachow-Moeller, N. and M. Svarer (2003). Estimation of the multinomial logit model with random effects. *Applied Economics Letters 10*, 389—392.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics.* New York: Academic Press.

Rabe-Hesketh, S., A. Pickels, and A. Skrondal (2001). *GLLAMM Manual, Technical Report.* London: Department of Biostatistics and Computing Institute of Psychiatry, Kings‘s College.

Revelt, D. and K. Train (1996). Incentives for appliance efficiency: Random-parameters logit models of household's models. *Manuscript, Department of Economics, University of California, Berkeley.*

Steiner, V. (2001). Unemployment persistence in the west german labour market: Duration dependence or sorting? *Oxford Bulletin of Economics and Statistics 1*, 91–113.

Steiner, V. and K. Wrohlich (2004). Household taxation, income splitting and labor supply incentives. a microsimulation study for germany. *CESifo Economic Studies 50*, 541–568.

Train, K. (2003). *Discrete Choice Models using Simulation*. Cambridge, UK: Cambridge University Press.

van Soest, A. (1995). Structural models of family labor supply: A discrete choice approach. *Journal of Human Resources 30*, 63–88.

Table 1: Estimation Results

	Conditional Logit		**Random Coefficient parametric**		**Random Coefficient nonparametric**	
	Coef.	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
income	-8.570	-1.940	-16.646	-2.690	-9.972	-1.550
income2	1.240	4.830	1.672	4.640	1.404	3.650
income*lm	-0.963	-3.060	-0.439	-1.190	-0.620	-1.770
income*lf	-0.602	-1.870	-0.480	-1.360	-0.828	-2.320
lm	59.225	10.750	57.203	9.590	64.272	10.560
lm2	-4.379	-13.100	-4.324	-12.690	-4.608	-13.210
lf	82.239	12.660	82.816	12.110	91.566	12.310
lf2	-7.154	-13.180	-7.104	-13.060	-7.426	-13.060
lf*lm	-1.986	-4.600	-2.183	-4.660	-2.872	-5.490
lm*ger	-1.072	-3.140	-1.093	-3.010	-1.086	-2.920
lf*ger	-0.218	-0.610	-0.240	-0.640	-0.202	-0.520
lm*lf*ger	-0.102	-0.800	-0.112	-0.830	-0.144	-1.060
income*ger	7.896	2.330	14.235	2.780	12.124	2.260
income2*ger	-0.590	-2.290	-1.001	-2.750	-0.865	-2.270
lm*east	-11.517	-4.840	-10.709	-4.200	-12.087	-4.600
lf*east	-13.334	-6.010	-12.587	-5.280	-13.979	-5.680
lm*lf*east	2.646	4.530	2.441	3.910	2.763	4.300
income*east	4.095	2.390	-0.210	-0.060	3.073	1.100
income2*east	-0.365	-2.650	-0.081	-0.330	-0.299	-1.440
lm*age	-0.396	-5.690	-0.456	-5.910	-0.480	-6.110
lm*age2	0.518	6.820	0.590	6.940	0.620	7.150
lf*age	-0.616	-6.810	-0.656	-6.850	-0.692	-6.850
lf*age2	0.843	8.040	0.895	8.040	0.946	8.030
lm*disabled	2.100	4.340	2.384	4.150	2.493	3.970
lf*disabled	2.830	3.580	3.057	3.630	3.078	3.470
lz*child6	4.215	15.690	4.331	15.390	4.491	14.770
lz*child16	2.136	11.160	2.150	10.780	2.203	10.470
lz*child17	0.512	2.740	0.542	2.770	0.543	2.660
d2	-1.051	-7.180	-0.967	-6.490	-1.004	-6.720
d11	-0.982	-12.130	-0.980	-12.080	-0.960	-11.790
d12	-0.492	-5.690	-0.491	-5.680	-0.463	-5.300
d16	-1.208	-11.480	-1.232	-11.690	-1.231	-11.660
d17	-0.551	-5.400	-0.552	-5.400	-0.554	-5.430
sd (income)	-	-	1.508	4.780	-	-
Var (income)	-	-	2.275	0.951	-	-
c1	-	-	-	-	-2.072	-4.730
c2	-	-	-	-	3.025	-
log p(c1)	-	-	-	-	0.378	0.850
p(c1)	-	-	-	-	0.5935	-
p(c2)	-	-	-	-	0.4065	-
Log-Likelihhod	-6044.168		-6038.912		-6032.448	
Akaike Criterion	4.3216		4.3183		4.3146	

Note: In the parametric estimation 10 adaptive quadrature points have been used. The non parametric distribution is described with 2 mass points. Log odds of the probabilities are estimated. The second mass point and its probability is calculated following the assumptions $E(c) = \sum_{i=1}^{N} \sum_{m=1}^{M} P_i(c^m)c^m = 0$ and $\sum_{m=1}^{M} P_i(c^m) = 1$. Variables: Income and leisure terms (lm, lf) are in logarithms. East and ger are dummy variables for East-Germany and German nationality. Dummy variables d2-d17 =1 for part time work. The sample consists of 2812 households drawn from the SOEP.

Table 2: Labor Supply Elasticities

	Male wage +1%			*Female wage +1%*		
	(1)	(2)	(3)	(1)	(2)	(3)
Change in participation rates (in percentage points)						
all	0.14	0.13	0.12	0.13	0.14	0.13
	(0.11 0.17)			*(0.11 0.15)*		
West Germany	0.15	0.14	0.13	0.15	0.16	0.15
	(0.13 0.18)			*(0.12 0.17)*		
East Germany	0.09	0.09	0.07	0.07	0.08	0.07
	(0.04 0.14)			*(0.03 0.10)*		
Change in hours (in percent)						
all	0.22	0.2	0.19	0.34	0.39	0.39
	(0.18 0.26)			*(0.28 0.40)*		
West Germany	0.24	0.22	0.21	0.39	0.45	0.45
	(0.20 0.27)			*(0.33 0.46)*		
East Germany	0.14	0.14	0.11	0.16	0.2	0.19
	(0.07 0.21)			*(0.07 0.25)*		

Note: (1) clogit, (2) parametric random coefficient, (3) parametric random coefficient. Numbers in parentheses are 95% bootstrap-confidence intervals (percentile method) based on 1,000 replications, which are derived from the conditional logit estimation. Source: SOEP, own calculations.