Laisney, François; Lechner, Michael

Working Paper

Combining panel data and macro information: an application to the estimation of a participation model

ZEW Discussion Papers, No. 93-23

Provided in cooperation with:
Zentrum für Europäische Wirtschaftsforschung (ZEW)

Discussion Paper

Discussion Paper No 93-23

Combining Panel Data and Macro Information: An Application to the Estimation of a Participation Model

François Laisney
Michael Lechner

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Labour Economics and Human Resources Series

Combining Panel Data and Macro Information: An Application to the Estimation of a Participation Model

by

François Laisney* and Michael Lechner**[1]

* *BETA, Université de Strasbourg and ZEW, Mannheim*

** *Institut für Volkswirtschaftslehre und Statistik, Universität Mannheim and ZEW, Mannheim*

November 1993

Abstract

When studying particular subgroups of a population, like for instance lone parents, the econometrician typically has few observations at hand. In such a situation, it is vital to take advantage of any valid complementary information that may be available. In this paper we illustrate, for the estimation of a participation model for lone mothers on data from the German Socio-Economic Panel 1984-1990, the relative benefits derived from using the panel structure of the data and from including macro information in the form of extra moments, as proposed by Imbens and Lancaster. The efficiency gains we find amount to having up to six times as many observations, and are shared almost equally between using the panel structure optimally and including macro information.

1 Support from the Deutsche Forschungsgemeinschaft (DFG) and from the European Commission under SPES grant CT910051 is gratefully acknowledged. Furthermore, we thank the DIW Berlin for providing us with the data of the Socio-economic panel (SOEP), and Susanne Salomon and Stefan Vögele for able research assistance. Finally, we benefited from discussions with Matthias Staat and Gerhard Wagenhals, our coauthors on the companion paper Laisney et al. (1993), and from the comments of participants in seminars and conferences in London, Mannheim, and Venice.

1 Introduction

When studying particular subgroups of a population, like for instance lone parents, the econometrician typically has few observations at hand. We found ourselves in such a situation when attempting to estimate participation and even labour supply models for lone mothers in West Germany on the basis of the only widely available data set containing relevant information, namely the Socio-Economic Panel (SOEP). In spite of many advantages, the SOEP is not ideal for studying such a special group as lone parents, because the number of such individuals in any single wave oscillates between 157 (1985) and 85 (1990). The obvious alternative would be to use a much larger sample to start with, and the obvious candidate would be the Mikrozensus, a 1% representative sample which is used, among other things, as a basis for the German Labour Force Survey. Unfortunately, neither the latter nor the original Mikrozensus is released by the Statistisches Bundesamt for public use (yet). If either were, one could even think of combining the informations it contains with those contained in the SOEP, as proposed for instance by Arellano and Meghir (1992).

On the other hand, the Statistisches Bundesamt publishes information on the basis of the Mikrozensus, and in our situation, it appears vital to take advantage of any valid complementary information we can obtain from that source.

The main purpose of this paper is to document the relative benefits derived from using the panel structure of the data and from including macro information in the form of extra moments, as proposed by Imbens and Lancaster (1991). This is widely applicable and extremely flexible. For empirical demand analysis, for instance, it can be contrasted with approaches that necessitate exact aggregation of some form in order to obtain identification or efficiency gains in the estimation of price reactions from the combined use of cross-section and macro information. Such approaches put artificial, and mostly empirically rejected, restrictions on the class of admissible micro models, since aggregation must result in a macro model that mimicks in some ways the micro model (see for instance Jorgenson et al., 1982, or Nichèle and Robin, 1993).

By contrast the approach followed here places no constraint, once identification is achieved, on the nature and number of extra moments used, provided that they are compatible with the micro information (and of course that some regularity conditions hold). In particular, there will be no necessity for the macro information to be be available at all dates corresponding to the waves of the panel, and it will be possible to discard some moments if their number makes computations problematic.

Moreover, the compatibility between the micro and the macro information can be tested before estimation is carried out, and this step enriches the data analysis that one should anyway perform before engaging in estimation.

The efficiency gains we find amount to having up to six times as many observations, and are shared almost equally between using the panel structure optimally and including macro information. However, another important by-product of this approach is, as alluded to above, that it disciplines the investigator into looking at the comparability of his sample information with that contained in published sources. A companion paper, Laisney et al. (1993) documents the data analysis thoroughly and puts the emphasis on the substantive results, also presenting the results of policy simulations.

Here we will therefore concentrate on the econometric aspects of the study: Section 2 describes the main features of the approach combining GMM estimation of limited dependent variable models on panel data and the use of extra moments retracing macro information; Section 3 presents the labour supply models used in the estimation; the data is reviewed shortly in Section 4 in order to allow the discussion of further aspects of the estimation strategy in Section 5; the results are presented in Section 6, and Section 7 gives a few concluding comments.

2 Main Econometric Aspects

Our approach combines the ideas concerning the estimation of limited dependent variable models on panel by the Generalised Method of Moments (GMM), outlined in Avery et al. (1983) and Breitung and Lechner (1993), with an approach combining micro and macro data sources in order to achieve better efficiency, suggested by Imbens and Lancaster (1991).

Using the panel structure means taking account of the fact that observations of the same individual over time may well be correlated. Moreover, we have to use an unbalanced panel, since otherwise we are left with too few observations to conduct any reasonable analysis (with our data, using a balanced panel means using observations on 20 individuals only). Increasing the efficiency is important given our small sample size. One way to attain this is to use full information maximum likelihood. Yet this is in general infeasible for $T = 7$, unless it is combined with restrictive assumptions on the covariance structure of the error terms. A more appealing approach is to use a method of moments where a first set of moments is given by the scores of the cross-section likelihood functions, and an additional set of moments takes advantage of the information provided by a very large dataset like the Mikrozensus. This is a 1% representative sample of total population, with over 600,000 individuals, which is large in comparison with the 15,000 individuals in the SOEP (we will ignore the sampling error there, and treat this information as giving exact knowledge of population parameters).

In the following we consider the vector of random variables (y, x) with $y = (y_1, \dots, y_t, \dots, y_T)'$ and $x = (x_1, \dots, x_t, \dots, x_T)'$, and the vectors (y_i, x_i), $i = 1, \dots, N$, of realisations from N independent draws from their joint distribution. The conditional distribution of y_t given x_t is assumed parametric, characterised by a parameter vector θ.

Given the moment restriction $E\Psi(y,x;\theta^0)=0$ and a weighting matrix C, the GMM estimator $\hat{\theta}$ of θ^0, the true parameter vector is defined as:

$$\hat{\theta} = \underset{\theta\in\Theta}{\operatorname{argmin}} \left[\frac{1}{N}\sum_{i=1}^{N} \Psi(y_i, x_i;\theta) \right]' C \left[\frac{1}{N}\sum_{i=1}^{N} \Psi(y_i, x_i;\theta) \right] \quad .$$

The optimal choice of the weighting matrix C is $\{E[\Psi(y,x;\theta^0)\Psi(y,x;\theta^0)']\}^{-1}$ or a sequence of random matrices converging to this expectation. Under some regularity conditions hold, the asymptotic distribution of the optimal GMM estimator for the case of independent observations is:

$$\sqrt{N}(\hat{\theta}-\theta^0) \overset{d}{\rightarrow} N(0,V) \quad , \text{ with}$$

$$V^{-1} = E\frac{\partial\Psi}{\partial\theta'}(y,x;\theta^0)'\{E[\Psi(y,x;\theta^0)\Psi(y,x;\theta^0)']\}^{-1} E\frac{\partial\Psi}{\partial\theta'}(y,x;\theta^0).$$

The covariance matrix can be consistently estimated by replacing expectations with sample averages, and θ^0 with a consistent estimate.[2]

Before proceeding we assume that attrition, the observability rule for the unbalanced panel, is ignorable, i.e. $E[\Psi_t(y_t,x_t;\theta^0) \mid r_t=1] = E[\Psi_t(y_t,x_t;\theta^0) \mid r_t=0]=0$. where r_{ti} equals one if the individual is observed in period and zero otherwise. This assumptions and using modified moments of the form $r_t\Psi_t(y_t,x_t;\theta^0)$ allows us to estimate all the moments necessary from the complete population without the need of further corrections. In order to facilitate the exposition and for the sake of brevity of notation, taking account of attrition will not be discussed any more.

Let us now partition the vector of moments according to $\Psi=(\Psi_1^{1\prime},\dots,\Psi_T^{1\prime},\Psi_1^{2\prime},\dots,\Psi_T^{2\prime})'$, and let $\Psi_t=(\Psi_t^{1\prime},\Psi_t^{2\prime})'$. This partition will correspond to the distinction between information from the panel data set used, and information obtained from macro data. The type of panel data models considered here allows for an arbitrary correlation structure over time, but requires that no component of the error term is correlated with any of the regressors. Under these assumptions, each cross section estimation using the marginal maximum likelihood estimator for (y_t,x_t) yields consistent parameter estimates of the components of θ^0 which can be identified from a single cross section. This suggests using the corresponding scores as elements of Ψ^1.

2 For a complete list of assumptions and proofs of the properties of the GMM estimator, see e.g. Hansen (1982).

It has first been noted by Avery et al. (1983) that furthermore imposing that the weighting matrix is the identity matrix, and that $\Psi(x, y; \theta) = \sum_{t=1}^{T} \Psi_t^1(y_t, x_t; \theta)$ yields the pooled estimator (i.e. the pseudo-maximum likelihood estimator obtained by ignoring the panel structure and treating all observations as if they were independent realisations). Hence the pooled estimator is consistent, as a GMM estimator. However, when computing its covariance matrix, the cross-period correlations between observations of the same individual have to be taken into account. We will use this pooled estimator as the benchmark from which to measure efficiency gains obtained either by using the panel structure optimally, or by drawing on macro information, or both.

The other estimators we shall consider use the scores of the marginal likelihoods as elements of Ψ^1. Thus they use information on more moments than when simply summing up the moments over time, as the pooling estimator does, and thereby are potentially more efficient. Moreover, they produce overidentifying restrictions which allow to test the specification of the model, in particular the constancy of coefficients over time.[3] By contrast, the approach of Avery et al. (1983) leads to an estimator which is exactly identified if only moments based on information of one period each are used. However, this potential efficiency gain over the pooled estimator comes at the cost of expanding the moment space: this may become a problem when estimating the optimal weighting matrix and the covariance matrix of the coefficient estimates.

The elements of Ψ^2 take outside information into account, as suggested by Imbens and Lancaster (1991). Since the outside information used here is based on a sample which is very large in comparison with our panel, it can be treated as representing the knowledge of population parameters. The details of the implementation (choice of moments, etc.) will be discussed below.

In order to conduct specification tests, we use the fact that under the null of a correct specification

$$\frac{1}{N}\sum_{i=1}^{N} \Psi(y_i, x_i; \hat{\theta})' \left[\sum_{i=1}^{N} \Psi(y_i, x_i; \hat{\theta}) \Psi(y_i, x_i; \hat{\theta})' \right]^{-1} \sum_{i=1}^{N} \Psi(y_i, x_i; \hat{\theta})$$

converges to a χ^2 distribution with a number of degrees of freedom equal to the rank of the covariance matrix of the moments minus the number of unrestricted parameters.

3 For a discussion of the relative efficiency of the various estimators that can be constructed along these lines, see Breitung and Lechner (1993).

3 A model for the labour supply of lone mothers

From the point of view of a lone mothers, the German tax and benefits system translates into a severely non convex budget set, mainly due to a one to one withdrawal rate on social benefits (SB). This results in marginal tax rates of 100% as long as earnings are below SB entitlements, and thus in a poverty trap: three fourths of the women in our sample would not be able to earn an income higher than their benefit entitlements as long as they work part time.[4]

Taking this into account within a model of continuous labour supply, as well as the endogeneity of wages related to the variability of the marginal tax rate, would result in a complicated model (see e.g. Hausman, 1985). Given our sample size, it seems unrealistic to hope to be able to estimate such a model. Following current practice, we therefore narrow the choices open to the individual down to three labour market states, namely non-participation, part-time and full-time work. This is supported by the existence of distinct spikes around 19 and 38 hours in the weekly hours distribution, corresponding to the two participation states mentioned.

Moreover, again bearing in mind that our sample size does not allow the estimation of complicated models, we follow Bingley et al. (1992) and Smith et al. (1991) in using the model proposed by van de Veen and van Praag (1981), which leads to the estimation of a simple bivariate probit.[5]

Let the three states $s = 0, 1, 2$ (non-participation, part-time and full-time work) be characterized by net income levels $y_s = wh_s + m_s - T_s$, where w denotes the gross wage, m_s is the unearned income in state s, which consists essentially of benefits in that state, B_s, and the tax liability, T_s, is given as $T_s = T(wh_s + m_s) - B_s$, where the tax function, T, takes all relevant aspects of the tax and benefit system into account. We assume that weekly hours of work take the values $h = \{0, 19, 38\}$ in the three states, respectively.

The preferences are supposed to be such that the decision to participate in the labour force ($P=1$) and the decision to work full time ($F=1$) can be described on the basis of a linear latent model:

$$P^*_{ti} = X_{ti}\gamma_t + u_{ti}, \quad P_{ti} = 1 \text{ if } P^*_{ti} > 0, \quad P_{ti} = 0 \text{ otherwise,}$$

$$F^*_{ti} = Z_{ti}\delta_t + v_{ti}, \quad F_{ti} = 1 \text{ if } P_{ti}F^*_{ti} > 0, \; F_{ti} = 0 \text{ otherwise.}$$

4 For a detailed description of the implications of the German tax and benefits system for lone mothers, see Laisney et al. (1993).

5 Here we will even end up concentrating on univariate probit estimation, for reasons that will become clear below.

where X contains socio-demographic characteristics as well as variables capturing effects of the availability of child care, in addition to the potential net income levels in the three different employment states; Z contains similar variables, but the income information for part time and full time work only.

If the error terms are identically normally distributed (and thus independent of the regressors), the conditional probabilities for the observed states of full-time work (FT), part-time work (PT) and non-participation (NP) given the regressors are:

$$P(NP_{ti}) = P(P_{ti} = 0) = 1 - \Phi\left(\frac{X_{ti}\gamma_t}{\sigma_{u_t}}\right) ,$$

$$P(PT_{ti}) = P(P_{ti} = 1 \wedge F_{ti} = 0) = \Phi\left(\frac{X_{ti}\gamma_t}{\sigma_{v_t}}\right) - \Phi_2\left(\frac{X_{ti}\gamma_t}{\sigma_{u_t}}, \frac{Z_{ti}\delta_t}{\sigma_{v_t}}, \rho_{u_t v_t}\right) ,$$

$$P(FT_{ti}) = P(F_{ti} = 1) = \Phi_2\left(\frac{X_{ti}\gamma_t}{\sigma_{u_t}}, \frac{Z_{ti}\delta_t}{\sigma_{v_t}}, \rho_{u_t v_t}\right) ,$$

where Φ, Φ_2 are the c.d.f. of the standardised univariate and bivariate normal distributions, and the standard deviations and the correlation coefficient are denoted by σ_{u_t}, σ_{v_t} and $\rho_{u_t v_t}$. Note that the probability for non-participation does not depend on the values of δ_t, $\rho_{u_t v_t}$, so that γ_t and $\sigma_{u_t}/\sigma_{u_0}$ can be consistently estimated by a (panel) binary participation probit.

However, a problem for the estimation of the model resides in the unobservability of the potential net incomes in the various employment states. In fact, all that is needed to compute these is the gross wage w. Predicting this requires the estimation of a wage equation. Again, here it would be infeasible to estimate the wage equation jointly with the participation and full-time equations, and we must resort to stepwise estimation, taking account (at least approximately) of potential selectivity problems, which is not as straightforward with panel data as it is with a single cross-section. Our choice for the wage equation is a selectivity corrected random effects regression with time dummies, based on human capital considerations. Identification hinges on the non-linearity of both the tax function and the functional form of the hazard.

Moreover, we face the problem that, due to the complexity of the tax function T, the exact reduced form for the participation equation, needed to compute the hazard, is not linear in w. Since many of the regressors we use are indicator variables, a possibility to cope with this would be to introduce interactions. But again, our small sample size leads to cells that are not sufficiently populated for meaningful computations. Thus we use the following short-cut: since the computation of the net income at zero hours does not require knowledge of w, we approximate the true reduced form with the participation equation with only y_0. The (possibly low) quality of the approximation will only affect the computation of the hazards.

For the purpose of this paper, which is mainly to illustrate the combined use of panel data and aggregate information, we concentrate on that fairly traditional reduced-form participation equation.[6]

4 Data

4.1 The dataset

The sample we use is an unbalanced panel drawn from the first seven waves of the socio economic panel of West Germany (SOEP, see Hanefeld, 1987, or Wagner et al., 1993, for an extensive description of this datasource). Our selection of observations of lone mothers is based on households classified in waves 1984 to 1990 as single parent households, with the lone mother being the head of the household. She is younger than 59, her oldest child living in the household is younger than 27 and her youngest child is younger than 21 years. The restrictions on the age of the children are enforced in order to exclude households with a second potential earner. After deleting cases with missing information in a particular wave, there are 296 individuals left. The means of the labour force states and of the explanatory variables are contained in table 1. More detailed information on the selection of the sample and more descriptive statistics can be found in Laisney et al. (1993).

4.2 Additional information

Recall that the outside information used here is based on a sample (*Mikrozensus*: the data is taken from the "Fachserien" of the Statistisches Bundesamt) which is more than 40 times larger than the SOEP: this allows us to ignore the sampling error in this information and to treat it as the knowledge of population parameters.

Due to data availability problems we use such information only for the participation equation. It consists in participation rates of lone mothers by age of the youngest child (younger than 15 or 18) and marital status (single, separated, divorced, widowed). These are available for 1985 to 1990. Decreasing the age for the youngest child resulted in all too sparse cells (given marital status). An additional moment for 1990 is constructed as the sum over exclusive age groups (25-34, 35-44, 45-55) of moments based on participation rates by age group and

6 Note however that, although the "structural" model outlined above allows the participation decision to be governed by a different process than the labour supply decision, it does not allow to recover the parameters of the underlying preference ordering - even leaving aside the identification problems connected with the fact that several of the regressors are also arguments of the tax function - so that the whole approach has a "reduced-form" character.

Table 1: Means of variables used in the estimation

Year	1984	1985	1986	1987	1988	1989	1990
dependent variable							
non participation	0.45	0.42	0.38	0.35	0.39	0.37	0.31
part time	0.12	0.13	0.15	0.18	0.15	0.22	0.26
full time	0.43	0.45	0.47	0.47	0.46	0.41	0.44
schooling							
Realschule	0.19	0.20	0.19	0.16	0.12	0.11	0.11
Abitur	0.05	0.06	0.07	0.09	0.11	0.14	0.19
number of children							
younger than 4	0.13	0.11	0.03	0.02	0.07	0.12	0.08
4-6 years old	0.25	0.15	0.19	0.14	0.10	0.07	0.09
younger than 7	0.37	0.25	0.22	0.16	0.15	0.19	0.16
7-14 years old	0.57	0.61	0.64	0.66	0.67	0.70	0.78
15-17 years old	0.35	0.32	0.34	0.31	0.35	0.32	0.35
*density of childcare * relevant child dummy*							
0-3 years / 10	0.03	0.04	0.00	0.00	0.02	0.03	0.01
4-6 years / 10	1.60	0.99	1.22	0.98	0.64	0.45	0.63
7-10 years / 10	0.05	0.06	0.07	0.07	0.07	0.10	0.09
age							
younger than 32	0.23	0.21	0.21	0.18	0.14	0.14	0.15
33-40	0.28	0.25	0.29	0.32	0.36	0.35	0.36
41-48	0.25	0.27	0.29	0.27	0.27	0.26	0.28
marital status							
single	0.11	0.11	0.12	0.13	0.08	0.11	0.16
divorced	0.45	0.47	0.43	0.45	0.45	0.47	0.58
widow	0.27	0.22	0.25	0.24	0.27	0.29	0.16
not german	0.23	0.19	0.19	0.19	0.22	0.23	0.19
regions							
northern	0.17	0.17	0.18	0.15	0.14	0.18	0.14
southern	0.30	0.26	0.26	0.33	0.33	0.30	0.33
urbanisation							
< 20'000	0.27	0.29	0.36	0.36	0.32	0.34	0.27
> 500'000	0.68	0.62	0.52	0.54	0.58	0.55	0.61
imputed income							
$Y^{NP}/10000$	1.69	1.78	1.80	1.82	1.82	1.88	1.87
$Y^{PT}/10000$	1.71	1.79	1.82	1.85	1.86	1.93	2.00
$Y^{FT}/10000$	2.02	2.04	2.09	2.15	2.32	2.37	2.66
observations	150	157	129	119	110	102	85

by marital status. Only the divorced with children younger than 18 years have been used, because the other cells do not contain enough observations. Furthermore, four coarser groupings have also been used, resulting in 24 additional moments. These are participation rates for lone parent females with (i) children younger than 6 years, (ii) age between 25 and 34, (iii) age between 35 and 44, and (iv) age between 45-54.

In the estimation, we use only the additional information that is compatible with the sample, and this can be read off Table 2, which also gives an overview on how the participation rates of specific subsamples of our data compare to the corresponding conditional participation rates in the population. Moreover, we discard the information on cells which contain less than ten observations in the sample.

The formula used to compute the test statistic for the compatibility of the macro and micro conditional employment probabilities is:

$$t_{jt} = \frac{N_{jt}(p_{jt} - \hat{p}_{jt})^2}{(1 - p_{jt})p_{jt}}$$

which is asymptotically distributed as $\chi^2(1)$. For wave t, the size of cell j is N_{jt}, $\hat{p}_{jt}$ is the sample frequency, whereas p_{jt} is its population (macro data) counterpart.

5 Further Considerations on Estimation

In order to avoid an excessive number of moments, those for the marital categories in each period have been added, resulting in only 12 additional moments instead of 48.

Using the expressions for the probablilities given in section 3, it is easy to derive the likelihood function for a single period, and given the usual regularity conditions it will be true that the expectation of the scores of these marginal likelihood functions will be zero for the true parameter values in each period. These scores will be used as elements of $\Psi_t^1(\cdot)$. Furthermore, recall that we have denoted p_j the participation probabilities in each particular socio-demographic group $j = 1, \ldots, J$, denoted by $\bar{p}_j$. Let j_{ti} be one if individual i belongs to group j in period t, and zero otherwise. The following expressions are used as elements of $\Phi_t^2(\cdot)$:

$$\Phi_t^2(\cdot) = \frac{1}{N}\sum_{i=1}^{N} r_{ti} j_{ti} \left\{ p_{jt} - \Phi\left(\frac{X_{ti}\gamma_t}{\sigma_{u_t}} \right) \right\}.$$

Two comments are in order here. Firstly, we have not used moments concerning the explanatory variables alone, although these can increase efficiency through their correlation with the other moments. The reason for this neglect is that, as mentioned above, we are concerned with the numerical problems that arise when the number of moments becomes important. Secondly, Imbens and Lancaster (1991) have shown that, given a partition based on explanatory variables, using marginal probabilities on top of the corresponding conditional probabilities is not informative, which justifies our disregard of such marginal probabilities.[7]

Since identification (in both equations, although here we focus on the participation equation) is only up to scale, at least one variance in both equations has to be normalised.[8] We have however chosen to normalise all variances, setting them to one, and restricted all coefficients, except the intercepts, to be constant over time. Correlations of the error terms over time are unrestricted.

Finally, all computations have been carried out on 30MHz 486 PCs using GAUSS. Some of the computations have been very time consuming, the extreme being three days for the joint estimation of the gross wage equation and the labour supply equation, none of which is reported here.

7 Strictly speaking, this statement needs to be qualified: we have not always used all conditional probabilities corresponding to a partition.

8 If none of the coefficients of an equation were unrestricted over time, all variances have to be normalised for that equation.

Table 2: Macro and micro information, and compatibility[9]

	1985				1986				1987				1988				1989				1990			
sample	#	M	m	t_i	#	M	m	t_i	#	M	m	t_i	#	M	m	t_i	#	M	m	t_i	#	M	m	t_i
separated																								
yc<15	25	0.48	0.52	0.14	22	0.52	0.59	0.14	14	0.50	0.79	4.48	12	0.52	0.50	0.01	9	0.51	0.56	0.07	6	0.56	0.67	0.30
yc<18	28	0.54	0.54	0.00	23	0.55	0.57	0.03	15	0.54	0.80	4.23	15	0.54	0.60	0.24	12	0.53	0.58	0.14	8	0.57	0.63	0.10
single																								
yc<18	15	0.56	0.53	0.05	15	0.57	0.67	0.58	15	0.56	0.67	0.76	8	0.57	0.63	0.09	11	0.56	0.64	0.24	14	0.59	0.71	0.85
divorced																								
yc<15	40	0.59	0.63	0.19	32	0.59	0.66	0.58	32	0.60	0.56	0.21	27	0.59	0.70	1.54	29	0.61	0.66	0.27	30	0.64	0.70	0.41
yc<18	55	0.65	0.65	0.02	42	0.64	0.71	1.11	43	0.64	0.65	0.02	41	0.63	0.71	0.95	37	0.64	0.70	0.61	41	0.68	0.73	0.60
widow																								
yc<15	12	0.44	0.33	0.52	11	0.41	0.27	0.82	9	0.42	0.22	1.46	12	0.42	0.50	0.28	13	0.45	0.46	0.01	4	0.54	0.75	0.72
yc<18	22	0.46	0.41	0.20	24	0.44	0.38	0.46	19	0.45	0.42	0.05	19	0.46	0.42	0.13	20	0.46	0.45	0.01	11	0.56	0.64	0.25
age																								
25-34	35	0.55	0.00	2.05	22	0.57	0.59	0.03	21	0.56	0.43	1.52	14	0.55	0.57	0.03	17	0.55	0.65	0.62	14	0.58	0.71	1.02
35-44	62	0.73	0.43	1.35	56	0.70	0.66	0.41	51	0.72	0.73	0.02	56	0.73	0.64	2.21	49	0.72	0.67	0.46	49	0.75	0.76	0.02
45-54	48	0.62	0.66	0.90	35	0.61	0.66	0.37	37	0.61	0.62	0.01	32	0.66	0.59	0.57	29	0.66	0.69	0.72	17	0.71	0.65	0.35
yc 0-6	34	0.43	0.32	1.57	26	0.45	0.42	0.10	19	0.47	0.53	0.22	17	0.50	0.41	0.53	19	0.46	0.37	0.57	15	0.49	0.47	0.03

1990, age mother	25-34				35-44				45-55			
	#	M	m	t_i	#	M	m	t_i	#	M	m	t_i
divorced, yc < 18	10	0.61	0.90	3.58	22	0.74	0.73	0.02	9	0.71	0.56	0.98

M = Macro, m = micro, yc = youngest child
t_i is $\chi^2(1)$.

9 This breakdown is not available for 1984.

6 Results

Tables 3 and 4 show the estimated coefficients and t-values for a single specification of the "reduced-form" participation equation but seven GMM estimators differing through the choice of the weighting matrix C and through whether or not macro information was used.[10] The reason why no macro information was used in combination with the pooled estimator is twofold. Firstly, recall that this pooled estimator is here simply a benchmark: it is the simplest consistent estimator at hand. Secondly, combining macro information with that estimator would have required a special and somewhat artificial programming effort: obviously, if one is prepared to engage in GMM estimation instead of simple probit, one will also go all the way to the optimal GMM estimator. This indeed, combined with macro information, constitutes the other extreme.[11]

Between these extremes, we report on estimates obtained with simpler choices for the weighting matrix. The first of these corresponds to the choice of the identity matrix, and it leads to surprisingly imprecise estimates in comparison with the simple pooling with time dummies. Our next choice is a diagonal matrix based on an estimate of the diagonal of $V[\Psi(y, x; \theta^0)]$. The consistent estimate of θ^0 needed for the estimation of this variance is obtained from the GMM estimation with the identity matrix. Again it comes as a surprise that the "diagonal" GMM estimates appear to be even less precise than the "identity" estimates, and also that in these two cases there appears to be no efficiency gain from the use of macro information (Table 5 gives the estimated standard errors of all estimates and is the best source for efficiency comparisons). In fact these results are counterintuitive only at first sight: since the computation of the estimated precisions is carried out at different estimates for θ^0, the theoretical asymptotic rankings need not apply.

Finally the "optimal" GMM estimator is computed on the basis of the full optimal weghting matrix, i.e. the inverse of $V[\Psi(y, x; \theta^0)]$, estimated at the "diagonal" GMM estimate.[12] Here at last we do find substantial efficiency gains. Comparing the first and third blocks of columns in table 4, i.e. the pooled with the optimal GMM including macro information, we see that the number of t-values above 3,

10 This is the same in all cases. Clearly there would have been scope also for comparison of different types of such information. Here it will suffice to say that in the first version of the study we had used less macro information, and found accordingly smaller efficiency gains.

11 The covariance matrix computed for the pooled estimator takes account of the correlations implied by the panel structure of the data (see Avery et al., 1983).

12 Because of the imprecision of both the "identity" and "diagonal" GMM estimates, we also did the computations on the basis of the pooled estimate. To our surprise, this had virtually no effect on the results. Thus we chose to report the results of what appears as a logical seqential procedure, where each step starts from the results obtained in the previous step.

leaving the intercepts aside, moves from 3 to 10. The most important change, at lest quantitatively, occurs for the variable "number of children below 4". For that variable, Table 5 shows that the ratio of the estimates of the asymptotic variances almost attains 6. In other words, the efficiency gain for this coefficient amounts to multiplying the sample size by 6. This is far from being as spectacular as the multiplier of 50 reported by Imbens and Lancaster for a very parsimonious participation equation for Dutch males, but it is still worth having. Part of the explanation of the difference between the efficiency gains in their study and ours is that the participation rate in their sample is above 90%, which means that the sample gives very little information about the parameters of interest. By contrast, the participation rates in our sample are between 45% and 35% (except for 1990, see Table 1), making the dichotomous variable much more informative in our case.

It is interesting to distinguish the efficiency gains realised through the optimal exploitation of the panel structure from the efficiency obtained from exploitation of the macro information. Comparing the first two groups of columns in Table 4, we see that the number of t-values above 3 went up to 6 only, and from Table 5 we see that the maximum ratio of estimated variances amounts to 2.5, so that the efficiency gains are approximately shared equally between the two sources. Note also the qualitative changes in the results when comparing the optimal GMM estimates with and without macro information: the negative impact of 7 to 14 years old children and of the age categories 33-40 and 41-48 on the paticipation probability becomes significant, whereas the negative impact of the net income at zero hours becomes insignificant.

In Table 5, the striking feature is that we find the expected decrease in standard errors for all coefficients when moving from pooled to optimal GMM without macro information and to optimal GMM with macro information (recall however that this is not automatically fulfilled, since the variance matrices are evaluated at different parameter vectors), whereas there are 25 order reversals in the comparisons between precisions in the absence and presence of macro information for the "identity" GMM, and 16 for the "diagonal" GMM. Moreover, the "diagonal" GMM does not perform uniformly better than the "identity" GMM, while both are uniformly (and severely) outperformed by the pooled estimator.

Appendix A reports similar results for the structural participation model outlined above. These lead to rigorously the same conclusions as the results discussed above as regards the efficiency gains. As regards the economic implications of the estimate, note the complete reversal in the signs of the income variables when moving from the pooled to the optimal GMM estimates. The latter have the right signs, although none is significant. The companion paper Laisney et al. (1993) reports on policy simulations on the base of the estimates in the last column of Table A.3.

Table 3: Estimates for the reduced-form participation equation - part 1

Weighting matrix	pooled		identity		identity		diagonal		diagonal	
Macroinformat.	no		no		yes		no		yes	
Variable	coef	t-val	coef	t-val	coef	t-val	coef	t-val	coef	t-val
time effects										
const 1984	1.20	2.5	1.71	1.4	2.87	1.9	1.61	1.3	1.43	1.1
const 1985	1.22	2.5	1.73	1.4	3.04	2.0	1.61	1.4	1.39	1.1
const 1986	1.33	2.6	1.87	1.5	3.12	2.1	1.72	1.5	1.51	1.2
const 1987	1.33	2.6	1.91	1.5	3.16	2.1	1.78	1.5	1.56	1.2
const 1988	1.27	2.5	1.86	1.5	3.11	2.0	1.72	1.4	1.49	1.2
const 1989	1.41	2.7	1.98	1.5	3.25	2.1	1.85	1.5	1.61	1.3
const 1990	1.48	2.7	2.05	1.6	3.35	2.1	1.89	1.6	1.67	1.3
schooling										
Realschule	0.52	2.2	0.43	1.5	0.41	1.4	0.41	0.8	0.46	0.8
Abitur	0.51	1.4	0.12	0.3	0.07	0.1	0.28	0.6	0.32	0.7
number of children										
younger than 4	-1.48	-4.1	-1.04	-1.8	-0.84	-1.7	-1.25	-2.0	-1.20	-1.9
4-6 years old	-0.20	-0.3	-1.21	-0.4	-2.25	-0.6	-1.03	-0.5	-1.02	-0.5
7-14 years old	-0.12	-1.0	-0.04	-0.2	-0.08	-0.4	0.03	0.1	-0.16	-0.6
15-17 years old	0.18	1.6	0.31	2.0	0.23	1.5	0.47	1.4	0.32	1.1
*density of childcare * relevant child dummy*										
0-3 years / 10	0.87	2.2	0.34	0.5	-0.94	-0.9	-0.05	-0.1	-0.28	-0.5
4-6 years / 10	-0.06	-0.5	0.11	0.2	0.25	0.4	0.08	0.2	0.06	0.2
age										
< 32	-0.16	-0.6	-0.37	-0.7	-0.41	-0.6	-0.20	-0.2	-0.29	-0.4
33-40	0.20	1.0	-0.03	-0.1	0.08	0.2	-0.18	-0.3	-0.41	-0.6
41-48	0.10	0.5	-0.07	-0.2	0.00	0.0	-0.40	-0.8	-0.74	-1.3
marital status										
single	0.04	0.1	0.09	0.1	-0.12	-0.2	-0.13	-0.2	-0.16	-0.2
divorced	-0.02	-0.1	0.08	0.2	-0.13	-0.3	-0.03	-0.1	-0.02	-0.0
widow	-0.81	-3.2	-0.74	-1.9	-0.82	-1.6	-0.84	-1.7	-0.89	-1.6
not german	0.51	2.7	0.14	0.5	0.13	0.3	-0.08	-0.2	-0.04	-0.1
regions										
northern	0.05	0.2	-0.04	-0.2	-0.15	-0.6	0.20	0.6	0.23	0.6
southern	0.57	3.0	0.38	1.5	0.25	0.9	0.50	1.2	0.42	1.0
urbanisation										
< 20'000	-0.23	-1.1	-0.39	-1.0	-0.49	-1.1	-0.47	-0.7	-0.60	-0.9
> 500'000	-0.33	-1.6	-0.52	-1.3	-0.64	-1.3	-0.41	-0.6	-0.47	-0.6
net income at 0										
$Y^{NP}/10000$	-0.43	-1.8	-0.55	-1.0	-1.00	-1.7	-0.46	-1.0	-0.08	-0.2
specification			$\chi^2(df)$	p-%	$\chi^2(df)$	p-%	$\chi^2(df)$	p-%	$\chi^2(df)$	p-%
distance	-		171		256		176	0.1	259	0.0
df			120	0.2	157	0.0	120		157	

Table 4: Estimates for the reduced-form participation equation - part 2

Weighting matrix	pooled		optimal		optimal	
Macroinformation	no		no		yes	
Variable	coef	t-val	coef	t-val	coef	t-val
time effects						
const 1984	1.20	2.5	1.78	4.7	1.30	4.2
const 1985	1.22	2.5	1.75	4.6	1.30	4.2
const 1986	1.33	2.6	1.85	4.7	1.43	4.7
const 1987	1.33	2.6	1.91	4.9	1.46	4.7
const 1988	1.27	2.5	1.77	4.5	1.39	4.5
const 1989	1.41	2.7	1.91	4.9	1.51	4.8
const 1990	1.48	2.7	1.98	4.9	1.57	4.9
schooling						
Realschule	0.52	2.2	0.68	3.6	0.57	3.6
Abitur	0.51	1.4	0.80	2.5	0.33	1.6
number of children						
younger than 4	-1.48	-4.1	-1.73	-7.4	-1.27	-8.7
4-6 years old	-0.20	-0.3	-1.56	-2.6	-1.15	-2.6
7-14 years old	-0.12	-1.0	-0.07	-0.9	-0.23	-3.6
15-17 years old	0.18	1.6	0.31	4.1	0.28	4.8
*density of childcare * relevant child dummy*						
0-3 years / 10	0.87	2.2	-0.12	-0.3	-0.41	-1.6
4-6 years / 10	-0.06	-0.5	0.16	1.8	0.08	1.2
age						
younger than 32	-0.16	-0.6	-0.33	-1.8	-0.23	-1.6
33-40	0.20	1.0	0.08	0.5	-0.36	-3.0
41-48	0.10	0.5	-0.22	-1.7	-0.68	-5.6
marital status						
single	0.04	0.1	-0.04	-0.2	-0.19	-1.2
divorced	-0.02	-0.1	0.10	0.7	0.03	0.3
widow	-0.81	-3.2	-0.89	-4.2	-0.78	-5.0
not german	0.51	2.7	0.21	1.2	-0.04	-0.3
regions						
northern	0.05	0.2	0.21	1.1	0.24	1.6
southern	0.57	3.0	0.84	5.1	0.56	4.6
urbanisation						
< 20'000	-0.23	-1.1	-0.50	-2.6	-0.59	-4.3
> 500'000	-0.33	-1.6	-0.68	-3.8	-0.52	-3.7
net income at 0						
$Y^{NP}/10000$	-0.43	-1.8	-0.53	-3.1	-0.05	-0.4
specification			$\chi^2(df)$	p-%	$\chi^2(df)$	p-%
distance	-		165	0.1	258	0.0
df			120		157	

Table 5: Standard errors of estimates

Weighting matrix	pooled	ident	ident	diag	diag	opt	opt
Macroinformation	no	no	yes	no	yes	no	yes
time effects							
const 1984	0.48	1.23	1.50	1.19	1.29	0.38	0.31
const 1985	0.48	1.25	1.49	1.15	1.24	0.38	0.31
const 1986	0.50	1.25	1.50	1.19	1.25	0.39	0.31
const 1987	0.51	1.27	1.53	1.21	1.27	0.39	0.31
const 1988	0.51	1.28	1.54	1.22	1.28	0.39	0.31
const 1989	0.52	1.29	1.56	1.21	1.27	0.39	0.31
const 1990	0.54	1.30	1.57	1.20	1.27	0.40	0.32
schooling							
Realschule	0.24	0.29	0.29	0.49	0.56	0.19	0.15
Abitur	0.36	0.47	0.53	0.49	0.44	0.32	0.21
number of children							
younger than 4	0.36	0.57	0.49	0.63	0.62	0.23	0.15
4-6 years old	0.74	3.29	3.75	2.15	1.94	0.59	0.44
7-14 years old	0.12	0.17	0.19	0.29	0.26	0.08	0.07
15-17 years old	0.11	0.15	0.16	0.34	0.29	0.07	0.06
*density of childcare * relevant child dummy*							
0-3 years / 10	0.39	0.72	1.06	0.46	0.61	0.37	0.25
4-6 years / 10	0.12	0.50	0.56	0.32	0.29	0.09	0.07
age							
younger than 32	0.25	0.55	0.66	0.87	0.84	0.19	0.15
33-40	0.20	0.36	0.41	0.64	0.63	0.16	0.12
41-48	0.18	0.32	0.36	0.53	0.57	0.14	0.12
marital status							
single	0.27	0.63	0.69	0.62	0.66	0.22	0.16
divorced	0.19	0.32	0.44	0.39	0.44	0.14	0.10
widow	0.25	0.40	0.52	0.49	0.56	0.21	0.16
not german	0.19	0.30	0.38	0.37	0.36	0.18	0.14
regions							
northern	0.23	0.25	0.26	0.32	0.37	0.19	0.15
southern	0.19	0.25	0.27	0.41	0.41	0.16	0.12
urbanisation							
< 20'000	0.21	0.40	0.46	0.63	0.67	0.19	0.14
> 500'000	0.21	0.41	0.47	0.68	0.74	0.18	0.14
net income at 0							
$Y^{NP}/10000$	0.23	0.54	0.58	0.47	0.44	0.17	0.13

Conclusions

What have we learned by doing this exercise? First of all that it is feasible, although costly, both in programming and computer time, to use the approach proposed by Imbens and Lancaster (1991) with panel data.

It is also costly in data analysis time, but this has the positive effect of leading to a better documentation of the data than is usual in studies based on micro data only.

The approach is flexible, in that the investigator only has to use the moments he finds published, retaining only those that are compatible with the data. Again this yields a by-product as regards data analysis, and the reward of the hard work needed is in the efficiency gains obtained.

Of course the whole approach is inferior to what could be achieved by getting hold of the micro data from which the aggregate data has been computed, but the latter is publicly available, and at very low cost, which is often not the case for the former.

Finally, for our example, the efficiency gains achieved proved vital in providing economically sensible results where less efficient estimators had produced doubtful ones.

Appendix A: Estimates for the Structural Participation Model

Table A.1: Estimation results of the structural participation equation - part 1

Weighting matrix	pooled		identity		identity		diagonal		diagonal	
Macroinformation	no		no		yes		no		yes	
Variable	coef	t-val	coef	t-val	coef	t-val	coef	t-val	coef	t-val
time effects										
const 1984	1.47	2.0	2.54	1.3	7.69	2.5	1.55	0.9	1.65	0.9
const 1985	1.48	2.1	2.52	1.3	7.68	2.6	1.55	0.9	1.63	0.9
const 1986	1.59	2.1	2.70	1.3	7.99	2.6	1.69	0.9	1.78	1.0
const 1987	1.60	2.1	2.73	1.3	8.15	2.6	1.72	0.9	1.80	1.0
const 1988	1.58	1.9	2.82	1.2	8.69	2.5	1.67	0.9	1.78	0.9
const 1989	1.72	2.0	2.91	1.3	8.83	2.5	1.76	0.9	1.86	1.0
const 1990	1.89	2.0	3.17	1.3	9.75	2.5	1.75	0.9	1.91	0.9
schooling										
Realschule	0.60	2.1	0.69	1.4	1.65	2.0	0.51	1.0	0.53	1.1
Abitur	0.95	1.4	0.90	0.6	2.67	1.4	0.15	0.2	0.49	0.5
number of children										
younger than 4	-1.44	-3.9	-0.90	-1.5	-1.03	-1.3	-1.08	-1.8	-1.28	-2.2
4-6 years old	-0.20	-0.27	-1.40	-0.4	-0.23	-0.1	-0.21	-0.1	-0.18	-0.1
7-14 years old	-0.09	-0.69	0.04	0.2	0.35	1.2	-0.04	-0.2	-0.08	-0.3
15-17 years old	0.19	1.60	0.30	2.0	0.31	1.5	0.51	1.3	0.42	1.2
*density of childcare * relevant child dummy*										
0-3 years / 10	0.86	2.2	0.34	0.4	0.80	1.0	-0.10	-0.2	0.10	0.3
4-6 years / 10	-0.05	-0.4	0.15	0.3	-0.00	-0.0	-0.05	-0.1	-0.06	-0.2
age										
younger than 32	-0.24	-0.9	-0.29	-1.6	-1.06	-1.4	-0.38	-0.4	-0.44	-0.5
33-40	0.19	1.0	-0.14	-0.9	0.04	0.1	-0.53	-0.8	-0.58	-0.9
41-48	0.12	0.6	-0.62	-4.5	0.24	0.5	-0.78	-1.3	-0.90	-1.4
marital status										
single	0.07	0.3	0.10	0.2	-0.21	-0.2	-0.07	-0.1	-0.17	-0.2
divorced	-0.02	-0.1	0.05	0.2	-0.18	-0.4	0.17	0.4	0.00	0.0
widow	-0.81	-3.3	-0.79	-2.1	-0.87	-1.9	-0.75	-1.5	-0.88	-1.8
not german	0.50	2.5	0.11	0.3	0.00	0.0	-0.03	-0.1	-0.1	-0.3
regions										
northern	0.07	0.3	-0.04	-0.2	-0.17	-0.5	-0.13	-0.3	0.15	0.3
southern	0.58	3.1	0.37	1.5	0.29	0.9	0.29	-0.5	0.46	1.0
urbanisation										
< 20'000	-0.23	-1.1	-0.34	-1.2	-0.35	-0.8	-0.44	-0.7	-0.55	-0.9
> 500'000	-0.32	-1.5	-0.46	-1.5	-0.54	-1.3	-0.39	-0.6	-0.37	-0.6
imputed income										
$Y^{NP}/10000$	0.38	0.3	-0.80	-0.3	-5.06	-1.4	-1.16	-0.6	-0.58	-0.3
$Y^{PT}/10000$	-0.72	-0.6	0.72	0.3	5.67	1.3	0.71	0.3	0.34	0.2
$Y^{FT}/10000$	-0.23	-0.4	-0.86	-0.5	-4.10	-1.7	0.11	0.1	-0.02	-0.0
specification test			χ^2(df)	p-%	χ^2(df)	p-%	χ^2(df)	p%	χ^2(df)	p%
distance	-		183	0.2	257	0.0	190		263	0.0
df			132		169		132		132	

Table A.2: Estimation results of the structural participation equation - part 2

Weighting matrix	pooled		optimal		optimal	
Macroinformation	no		no		yes	
Variable	coef	t-val	coef	t-val	coef	t-val
time effects						
const 1984	1.47	2.0	1.31	2.6	1.43	4.1
const 1985	1.48	2.1	1.28	2.6	1.43	4.1
const 1986	1.59	2.1	1.43	2.8	1.59	4.6
const 1987	1.60	2.1	1.41	2.8	1.59	4.5
const 1988	1.58	1.9	1.19	2.1	1.55	4.1
const 1989	1.72	2.0	1.33	2.4	1.65	4.4
const 1990	1.89	2.0	1.27	2.1	1.68	4.1
schooling						
Realschule	0.60	2.1	0.71	3.5	0.60	3.8
Abitur	0.95	1.4	0.21	0.5	0.45	1.6
number of children						
younger than 4	-1.44	-3.9	-1.62	-7.3	-1.36	-9.1
4-6 years old	-0.20	-0.3	-0.37	-0.9	-0.23	-0.6
7-14 years old	-0.09	-0.7	-0.14	-1.7	-0.17	-2.7
15-17 years old	0.19	1.6	0.38	4.7	0.39	7.2
*density of childcare * relevant child dummy*						
0-3 years / 10	0.86	2.2	0.10	0.4	-0.00	-0.0
4-6 years / 10	-0.05	-0.4	-0.04	-0.7	-0.05	-0.9
age						
younger than 32	-0.24	-0.9	-0.29	-1.6	-0.40	-2.8
33-40	0.19	1.0	-0.14	-0.9	-0.56	-5.3
41-48	0.12	0.6	-0.62	-4.5	-0.86	-7.8
marital status						
single	0.07	0.3	-0.03	-0.1	-0.19	-1.3
divorced	-0.02	-0.1	0.21	1.6	0.03	0.4
widow	-0.81	-3.3	-0.75	-3.8	-0.80	-5.4
not german	0.50	2.5	0.23	1.3	-0.16	-1.2
regions						
northern	0.07	0.3	-0.05	-0.3	0.15	1.0
southern	0.58	3.1	0.69	4.3	0.58	5.1
urbanisation						
< 20'000	-0.23	-1.1	-0.51	-2.7	-0.51	-4.2
> 500'000	-0.32	-1.5	-0.67	-3.8	-0.39	-3.0
imputed income						
$Y^{NP}/10000$	0.38	0.3	-1.34	-2.5	-0.49	-1.3
$Y^{PT}/10000$	-0.72	-0.6	0.79	1.4	0.28	0.7
$Y^{FT}/10000$	-0.23	-0.4	0.37	1.1	0.05	0.2
specification test			$\chi^2(df)$	p%	$\chi^2(df)$	p%
distance	-		185	0.2	261	0.0
df			132		169	

Table A.3: Standard errors of estimates

Weighting matrix	pooled	ident	ident	diag	diag	opt	opt
Macroinformation	no	no	yes	no	yes	no	yes
time effects							
const 1984	0.73	2.02	3.03	1.79	1.79	0.50	0.35
const 1985	0.72	2.00	2.98	1.75	1.75	0.49	0.36
const 1986	0.74	2.07	3.07	1.80	1.81	0.50	0.35
const 1987	0.76	2.10	3.15	1.82	1.83	0.51	0.35
const 1988	0.84	2.33	3.47	1.96	1.95	0.57	0.38
const 1989	0.84	2.30	3.48	1.93	1.93	0.56	0.38
const 1990	0.96	2.55	3.91	2.06	2.05	0.61	0.41
schooling							
Realschule	0.28	0.49	0.81	0.49	0.49	0.20	0.16
Abitur	0.70	1.51	1.97	0.94	0.92	0.43	0.29
number of children							
younger than 4	0.37	0.61	0.77	0.60	0.57	0.22	0.15
4-6 years old	0.74	3.56	4.23	2.17	1.93	0.41	0.37
7-14 years old	0.13	0.23	0.30	0.27	0.25	0.08	0.06
15-17 years old	0.12	0.15	0.20	0.41	0.35	0.08	0.05
*density of childcare * relevant child dummy*							
0-3 years / 10	0.39	0.77	0.80	0.44	0.38	0.27	0.23
4-6 years / 10	0.11	0.54	0.64	0.33	0.30	0.06	0.06
age							
younger than 32	0.27	0.66	0.78	0.97	0.90	0.19	0.14
33-40	0.20	0.39	0.45	0.70	0.66	0.15	0.11
41-48	0.19	0.33	0.45	0.62	0.64	0.14	0.11
marital status							
single	0.27	0.67	0.88	0.74	0.74	0.22	0.15
divorced	0.19	0.31	0.42	0.39	0.42	0.13	0.10
widow	0.25	0.38	0.47	0.51	0.50	0.20	0.15
not german	0.20	0.32	0.47	0.41	0.38	0.18	0.14
regions							
northern	0.22	0.24	0.32	0.52	0.58	0.18	0.14
southern	0.19	0.25	0.33	0.48	0.46	0.16	0.11
urbanisation							
< 20'000	0.21	0.30	0.41	0.59	0.60	0.19	0.12
> 500'000	0.21	0.30	0.40	0.65	0.66	0.18	0.13
imputed income							
$Y^{NP}/10000$	1.30	2.93	4.29	2.30	2.06	0.64	0.47
$Y^{PT}/10000$	1.13	2.87	4.47	2.21	1.97	0.58	0.39
$Y^{FT}/10000$	0.61	1.64	2.39	1.01	0.96	0.34	0.22

References

Arellano, M. and C. Meghir (1992): "Female Labour Supply and On the Job Search: An Empirical Model Estimated Using Complementary Data Sources", *Review of Economic Studies*, 59, 537-557.

Avery, R.B., L.P. Hansen and V.J. Hotz (1983): "Multiperiod Probit Models and Orthogonality Condition Estimation", *International Economic Review*, 24, 21-35.

Bingley, P., E. Symons and I. Walker (1992): "The Labour Supply of UK Lone Mothers: the Effect of Maintenance, and the Welfare System", *IFS Working Paper W92/19*, Institute for Fiscal Studies, London.

Breitung, J. and M. Lechner (1993): "GMM estimation in the non-linear model with panel data", *mimeo*.

Hanefeld, U. (1987): "Das Sozioökonomische Panel", Berlin: Campus Verlag.

Hansen, L.P. (1982): "Large Sample Properties of Generalised Method of Moments Estimators", *Econometrica*, 50, 1029-1054.

Hausman, J. (1985): "The Econometrics of Nonlinear Budget Sets", *Econometrica*, 53, 1255-1282.

Imbens, G.W. and T. Lancaster (1991): "Combining Micro and Macro Data in Microeconometric Models", *Discussion paper 1578*, Harvard University.

Jorgenson, D.W., L.J. Lau and T.M. Stoker (1982): "The Transcendental Logarithmic Model of Aggregate Consumer Behavior", in *Advances in Econometrics*, R.Basman and G.Rhodes (eds.). Greenwich: JAI Press.

Laisney, F., M. Lechner, M. Staat and G. Wagenhals (1993): "Labour Force and Welfare Participation of Lone Mothers in West Germany", *ZEW Discussion paper 93-25*.

Lanot, G., E. Symons and I. Walker (1991): "The Labour Supply and Wages of UK Lone Mothers: a Discrete Choice Model", CREST-ENSAE, Paris, Institute for Fiscal Studies, London.

Nichèle, V. and J.M. Robin (1993): "Simulation of Indirect Tax Reforms Using Pooled Micro and Macro French Data", *Journal of Public Economics*, forthcoming.

Smith, N., I. Walker and N. Westergård-Nielsen (1991): "The Labour Supply of Danish Lone Mothers", *mimeo*, Århus University.

Van de Veen, W.P.M.M., and B.M.S. van Praag (1981): "The Demand for Deductibles in Private Health Insurance", *Journal of Econometrics*, 17, 219-223.

Wagner, G., R. Burkhauser, and F. Bheringer (1993): "The English Language Public Use File of the German Socio Economic Panel Study", *Journal of Human Ressources*, 28, 429-433.