Panel Regression Models for Measuring Poverty Dynamics in Great Britain

Gianni Betti, Antonella D'Agostino, Laura Neri

Dip. Metodi Quantitativi, University of Siena

Premise[1]

In this paper we aim to study poverty dynamics and the socio-demographic factors influencing it; alternative measures for the definition of the concept of poverty, making use of panel regression models, are compared. The introduction of two alternative measures is necessary in order to overcome some limitations of the so-called traditional approach.

In the traditional approach to poverty measurement, a statistical unit (individual or household) is defined as poor if its net equivalent household income is below a certain threshold, the poverty line, i.e. a percentage of the mean or median of the overall income distribution. Let us now consider some limitations.

The division of the population into the dichotomy of the poor and the non poor seems to be an over simplification; in fact, as pointed out by Cheli and Lemmi (1995) "...

[1]**This work was co-financed by Murst funds for the projects "Occupazione e disoccupazione in Italia: misura e analisi dei comportamenti", and was supported by the Training and Mobility for Researcher Access to Large Scale Facilities Programme, benefitted by the three outhors at the European Centre for Analysis in the Social Sciences (ECASS). The paper is the result of the common work of all the authors; in particular G. Betti has written sections 2, 4.1 and 4.3.1; A. D'Agostino has written sections 3, 4.2 and 4.4; L. Neri has written sections 1, 4.3.2 and 5.**

poverty is not a simple attribute that characterises an individual in terms of its presence or absence"; we also believe that the relative well-being of a statistical unit is a matter of degree. Moreover, poverty measures based only on monetary variables are not able to capture all the aspects and nuances of the phenomenon; multidimensional indicators seem to be more appropriate in order to describe deprivation better. Finally, in a dynamic context, the traditional approach has a further limitation: the mobility of the units near to the poverty line is overestimated.

Another aspect must be discussed is that, in the context of poverty dynamics analysis there is no unanimity in the choice of the longitudinal units; the controversy is about choosing individuals or households. In any case, it seems reasonable to consider variables related to the household as poverty indicators, particularly when a multidimensional concept of poverty is considered. On the other hand, it is difficult to define the household as a longitudinal unit of analysis in any rigorous way. The position of researchers on the definition of the unit of analysis in a longitudinal study can be found in, among others, Buck *et al.* (1995), Lillard and Willis (1978), Duncan *et al.* (1984), Bane and Ellwood (1986), Stevens (1995), Trivellato (1998).

Some advantages are ensured by following the individuals across time, and considering them as units of analysis: household and individual units remain representative of the population structure across time. If the individual is chosen as a unit of analysis, it is necessary to consider that the poverty definition relies on the household; for this reason, a household poverty indicator must be assigned to each household member.

In this work, the individual has been chosen as a unit of analysis, according to the BHPS survey design. The household poverty indicators relating to the two measures have been calculated and assigned to each household member aged 16+; in the same way household/head of household characteristics have been assigned to each household member time-by-time.

The paper is organized in the following way. In section 1, two alternative measures for the definition of the concept of poverty are presented. The panel regression models are presented in section 2: particular attention is paid to the treatment of the unobservable heterogeneity among individuals. The empirical analysis, reported in section 3, is based on the data set collected by the BHPS from 1991 to 1997; finally some concluding remarks end the paper in section 4.

1. Poverty definitions

We start from the original paper of Cerioli and Zani (1990) and from the Totally Fuzzy and Relative approach of Cheli and Lemmi (1995) in order to define multidimensional fuzzy poverty measures. The adoption of a multi-dimensional approach involves two main problems: the choice of the indicators and the aggregation process.

Although deprivation is widely recognised as a multidimensional phenomenon, we still believe that indicators based on monetary variables have a fundamental role and therefore are worthy of special treatment. For this reason two different fuzzy measures are considered: the first one is based on a monetary variable only and we refer to it as Fuzzy Monetary (FM); the second measure is based on several indicators related to housing conditions, durable goods, and so on and we refer to it as Fuzzy Supplementary (FS). Let us consider the construction of the two measures.

The monetary variable utilised for the FM method consists in the net equivalent household income z_{jt}; making use of the concepts of the fuzzy set theory (Zadeh, 1965;

Dubois and Prade, 1980), the degree of deprivation of any household j at any period t is defined as the membership function to the fuzzy set of poor:

$$\mu\left(z_{jt}\right)=\left[1-F\left(z_{jt}\right)\right]^{\alpha_t}=I_{jt}^{FM} \qquad j=1,...,J \;\; t=0,1,...,T \tag{1.1}$$

where $F(.)$ is the household distribution function according to the equivalent income. As proposed by Cheli (1995) we determine parameters α_t so that the membership function means are not merely equal to 0.5, but are equal to the proportion of poor units according to the traditional approach (the so called head count ratio H).

In order to identify the year-by-year household head count ratios H_t, poverty lines need to be calculated as in the traditional approach; this can be done by following two strategies: i) at each period the poverty line is half of the household equivalent income mean; ii) the poverty line is calculated for the first period only and is kept fixed (in real terms) for the following years; we adopt the latter strategy as in Jenkins (1999).

The FS measure is based on some supplementary variables x_{jtk} $(k=1,...,K)$, such as amenities in the household, ability to afford durable goods, accommodation problems, and any other variable relevant for the multidimensional definition of deprivation. The construction process of this measure is fully described in Betti and Verma (2000).

When supplementary variables are ordinal with two or more categories, for each variable k, with ordered categories 1 (least deprived) to M (most deprived), we define the single poverty indicator for all households in category m as follows:

$$s_{jtk}=\frac{m-1}{M-1} \tag{1.2}$$

When supplementary variables are quantitative poverty indicators they can be calculated in a way similar to equation (1). The aggregation process of the single indicators into the multidimensional measure is described by a weighted mean:

$$s_{jt} = \frac{\sum_{k=1}^{K} w_k \cdot s_{jtk}}{\sum_{k=1}^{K} w_k} = I_{jt}^{FS} \tag{1.3}$$

The weights w_k are determined by two statistical considerations: i) firstly, the weight is determined by the variable's power to "discriminate" among individuals in the population, that is, by its dispersion; ii) from a non-redundant point of view, it is necessary to limit the influence of those characteristics that are highly correlated to the others. For a detailed description of the weights see Betti and Verma (2000).

2. Poverty indicator model

The basic element of our approach to model poverty dynamics is a poverty indicator function.

Let $I_{it}^{(\bullet)} = \left(I_{it}^{FM}, I_{it}^{FS}\right)$; the two indicators represent respectively the poverty indicator based on income (FM measure) and the poverty indicator based on qualitative variables (FS measure) at time t ($t = 0, ..., T_i$) for each individual aged 16+ i ($i = 1, ..., N$), in household j ($j = 1, ..., J$) introduced in section 1.

As I_{it}^{FM} and I_{it}^{FS} range in the interval $[0 - 1]$, a logit transformation is performed in order to create two variables ranging between $-\infty$ and $+\infty$; we obtain:

$$
\begin{aligned}
y_{it}^{FM} &= \text{logit}(I_{it}^{FM}), \\
y_{it}^{FS} &= \text{logit}(I_{it}^{FS}).
\end{aligned}
\tag{2.1}
$$

The poverty indicator function for each indicator is assumed to be:

$$
y_{it}^{(\bullet)} = \beta^{'}\mathbf{x}_{it} + \Psi_t + u_{it}, \tag{2.2}
$$

where $\mathbf{x}_{it}$ is a vector of k time-varying exogenous variables observed on individual i representing the effect of observed heterogeneity, $\Psi_t = f^p(t)$ is a polynomial of grade p that represents the effect of time, u_{it} is the error structure and β is a vector of k unknown parameters.

The error structure is of the form:

$$
u_{it} = \delta_i + \xi_{it}, \tag{2.3}
$$

where ξ_{it} has a first-order autoregressive structure, e.g.

$$
\xi_{it} = \rho\xi_{it-1} + \eta_{it}. \tag{2.4}
$$

In (2.3) δ_i represents a random individual component distributed as $N(0, \sigma_\delta^2)$; in (2.4) η_{it} is a purely random component *i.i.d.* assumed to be distributed as $N\left(0, \sigma_\eta^2\right)$ and ρ is the serial correlation coefficient common to all individuals. The random variables δ_i and η_{it} are also assumed to be independent of each other and of $\mathbf{x}_{it}$ and Ψ_t[2].

[2] A similar structure has been introduced for modelling earnings mobility by Lillard and Willis (1978),

The specified model in (2.2) combines autoregression with variance components to obtain a model allowing for both heterogeneity and autocorrelation (Anderson and Hsiao, 1981; Mansour *et al.*, 1985; Goldstein *et al.*, 1994). The individual component of this error structure, δ, represents the effect of individual unobserved (or unobservable) heterogeneity in equation (2.2) and this effect is assumed to persist through out the period of observation.

The serial correlation term, ρ, represents the rate of deterioration of the effects of random shocks η persisting for more than one year; it may also reflect the effect unobserved individual variables serially correlated over time, i.e., with a slow change across time.

In econometric literature such a model is called *serial correlation model* as y_{it} is only affected by x_{it} not by x_{it-1} , in other words if $\mathbf{x}$ is increased in period t and then returned to its former level, the distribution of y in period $t+1$ is not affected. Past y is informative because it helps to predict the effect of unobservable variables which are serially correlated; this model also implies that y_{it} fluctuates around the equilibrium level $\left(\beta^{'}\mathbf{x}_{it} + \Psi_t + \delta_i\right)$ as do the effects of unobservable variables $\{\xi_{it}\}$ that follows a first-order autoregressive process.

As reported in Anderson and Hsiao (1982) contrary to the case of the dynamic model for a single time series, the assumption concerning the initial observations plays a crucial role in interpreting the model and in devising consistent estimates.

For this reason a special assumption is made regarding the distribution of the first response on each unit; this is taken to be the marginal distribution,

$$y_{i0} \sim N\left(\beta^{'}\mathbf{x}_{i0}, \frac{\sigma_\eta^2}{1-\rho^2}\right). \tag{2.5}$$

where the authors have specified a simple earnings function.

This is stationary and does not depend on previous values. For $t > 1$ we assume that the residual covariance structure is of the form

$$E\left(u_{it}u_{jt'}\right) = \begin{cases} \sigma_\delta^2 + \frac{\sigma_\eta^2}{1-\rho^2} & i = j \ t = t' \\ \sigma_\delta^2 + \rho^S \frac{\sigma_\eta^2}{1-\rho^2} & i = j \ |t - t'| = S > 0 \\ 0 & i \neq j \end{cases} \cdot \tag{2.6}$$

As we are dealing with unbalanced panel data due to missing observations, the covariance structure reported in (2.6) presents no rows and columns which correspond to the missing observations (Jones, 1993).

3. Empirical analysis

The empirical analysis has been conducted using the data set of the British Household Panel Survey from 1991 to 1997 (Waves 1 to 7). The BHPS is a complex panel survey representing a unique source of longitudinal information on incomes and other variables at household and individual level in Britain (Taylor, 1994; Taylor *et al.*, 1996; Taylor, 1998).

In section 3.1, cross-sectional household and individual poverty indicators are calculated according to the FM and FS measures; the dynamic models are presented in section 3.2, while parameter estimates are shown in section 3.3. A comparison between the two measures is reported in section 3.4.

3.1. Cross-sectional poverty indicators

The sample used to construct the household poverty indicators (see equations (1.1) and (1.3)) consists of those households in which all eligible adults gave a full interview (i.e.

Table 3.1: Household membership function means

Wave	1	2	3	4	5	6	7
$E\left[I_{jt}^{FM}\right] = H_t$	0.197	0.159	0.157	0.148	0.138	0.126	0.129
α_t	4.165	4.661	4.480	4.677	4.737	4.832	4.874
$E\left[I_{jt}^{FS}\right]$	0.418	0.394	0.372	0.354	0.338	0.321	0.304
J	4826	4556	4354	4378	4259	4372	4383

BHPS variable *wifho=10*); in this data set the net equivalent household income is present for all individuals (Bardasi *et al.*, 1999); missing values in the supplementary variables have been imputed using the approach adopted by Raghunathan *et al.* (1997).

The household distribution function $F(.)$ in equation (1.1) has been estimated parametrically[3] on the basis of the net equivalent household income from the 1991 data set consisting of 4826 households. For the same reference year, the poverty line has been calculated as half of the mean net equivalent household income; the line results as being equal to £ 135.45 per week deflated to January 1998 prices.

Table 3.1 reports the percentages of poor households in waves 1-7 according to the traditional approach (the head count ratios H_t) and the values of parameters α_t of formula (1.1) so that:

$$E\left[I_{jt}^{FM}\right] = E\left[1 - F\left(z_{jt}\right)\right]^{\alpha_t} = H_t \tag{3.1}$$

Therefore the head count ratios coincide with the household membership function means calculated year-by-year. These show a descending behavior from 1991 to 1996, while there is a slight increment in the final year.

In order to evaluate the household membership functions according to the FS mea-

[3] The model utilised is due to Dagum (1977); it has the form $F(z) = \left(1 + \lambda z^{-\delta}\right)^{-\beta}$; $z > 0$; $\beta, \lambda > 0$; $\delta > 1$; where β and δ are shape parameters, whereas λ is a scale parameter. The ML estimates for year 1991 are: $\widehat{\beta} = 0.6559$; $\widehat{\lambda} = 18.9579$; $\widehat{\delta} = 3.3566$.

sure (formula (1.3)) several supplementary variables are considered; they refer to housing conditions and to the presence of durable goods; the exhaustive list of poverty symtoms is: house which is not owned; lack of central heating, colour TV, videorecorder, washing machine, dishwasher, home computer, CD player, microwave, car or van.

We would like to underline that the indicators reported in the previous list are not proper poverty symptoms: sometimes, it could merely be a matter of choice whether to own a car or not (especially if someone lives in Central London); therefore it would be more informative to know whether someone can afford a particular good. Unfortunately, this information is not collected by the BHPS, at least in the first waves.

Let us now analyse household means of the FS indicators; they are reported in the fourth row of Table 3.1: in this case we can observe a regular decrement of the indicators across seven years.

Household indicators I_{jt}^{FM} and I_{jt}^{FS} are assigned to each individual in household j, in order to get individual membership functions I_{it}^{FM} and I_{it}^{FS}. The year-by-year averages of these indicators for all individuals are reported in Table 3.2, whereas the means calculated for the whole sample and disaggregated by regions, only on the basis of individuals aged 16+ are in Table 3.3; The temporal sequence of the means reflects that of households reported in Table 3.1.

According to the FM measure, we can observe that averages for individuals aged 16+ are constantly lower than those for all individuals; on the contrary, according to the FS measure, the averages for individuals aged 16+ are constantly higher than those for all individuals.

This seems to suggest preliminary and interesting evidence: although households with children are more deprived in terms of monetary variables, they show a higher standard

Table 3.2: Membership function means; all individuals

Wave	1	2	3	4	5	6	7
$E\left[I_{it}^{FM}\right]$	0.185	0.156	0.151	0.141	0.129	0.121	0.125
$E\left[I_{it}^{FS}\right]$	0.365	0.342	0.320	0.301	0.285	0.270	0.254
ALL	11634	11001	10475	10477	10128	10544	10555

of living according to the supplementary variables; this can be explained by the inclusion of variables such as possession of colour TV, videorecorder, CD player, home computer, i.e. durable goods highly diffused in families with children. However, it is important to underline that, although the FM measure has been obtained by discounting income for household typology, the FS measure does not take the scaling into account.

3.2. Model specification

The analysis refers to the unbalanced panel of individuals aged 16+ created from the household sample described in the previous section. The total sample size is of 12766 individuals and 57964 repeated measures.

The models specified in (2.2) have been estimated according to three macro-regions[4] the South-East, the West and the North[5]; in each model the dependent variable consists, alternatively, of one of the two poverty indicators.

The time indicator is the variable PEPI. The first is of models are characterized by a time dependence specified as a polynomial of degree $p = 1$; in other words a linear trend assumption has been made. Table 3.3 approximately shows a linear trend for the sample data, particularly evident for the FS indicator. Anyway, it is important to

[4] It is important to point out that this choice was been necessary for computational problems given the large sample size. However, this analysis has allowed us to make very interesting territorial comparisons.

[5] The macro region called South-East is composed of London (inner and outer), the South-East, East Anglia, East Midlands; the region called West is composed of the South West, Midlands, Manchester, Merseyside and Wales; the macro region called North is composed of Yorkshire, the Region of North, Yorks & Humber, Tyne &Wear and Scotland.

Table 3.3: Membership function means: individuals aged 16+, whole sample and regions

		1	2	3	4	5	6	7
Whole sample	$E\left[I_{it}^{FM}\right]$	0.172	0.141	0.137	0.127	0.118	0.109	0.112
	$E\left[I_{it}^{FS}\right]$	0.377	0.354	0.331	0.313	0.296	0.279	0.264
	N	8948	8497	8108	8125	7858	8199	8229
South-East	$E\left[I_{it}^{FM}\right]$	0.157	0.133	0.124	0.116	0.112	0.105	0.105
	$E\left[I_{it}^{FS}\right]$	0.361	0.342	0.318	0.305	0.290	0.271	0.258
	N	3567	3439	3316	3335	3283	3414	3455
West	$E\left[I_{it}^{FM}\right]$	0.183	0.144	0.147	0.140	0.120	0.109	0.107
	$E\left[I_{it}^{FS}\right]$	0.384	0.361	0.342	0.322	0.308	0.294	0.278
	N	2689	2494	2351	2367	2253	2369	2359
North	$E\left[I_{it}^{FM}\right]$	0.181	0.148	0.145	0.130	0.126	0.116	0.126
	$E\left[I_{it}^{FS}\right]$	0.397	0.368	0.347	0.328	0.309	0.291	0.278
	N	2683	2564	2441	2423	2322	2416	2415

underline that this assumption was made after different trials using polynomials of a greater degree or a non parametrically time dependence. In order to compare results of the parameter estimation they have been standardized (variable names are LGFAST for y_{it}^{FM} and LGFAQST for y_{it}^{FS}).

In the second group of models, a vector of k time-varying exogenous variables are introduced as well as the trend.

The variables considered in the analysis refer to household characteristics. The variables referring to the household head are: a dummy variable for the gender, SEX (1 if male); the age and the age square, AGE and AGE2; two dummies for the employment status, JBSTA1 (1 if self or in paid employment) and JBSTA2 (1 if unemployed); two dummies for school qualification, QUAL1 (1 if high degree: A level or higher qualification) and QUAL2 (1 if medium degree: O level, CSE level or other qualification), a dummy variable for the marital status, MASTA (1 if married or couple). The variable referring to the household are three dummies for the presence of children in the household: children aged 0-4, CH0_4 (1 if present), aged 5-11, CH5_11, (1 if present), aged 12-18, CH12_18

(1 if present); two specifications for the household size SIZE and SIZE2 (size square). All variables are time-dependent[6].

All models have been estimated by marginal likelihood estimation using the program MIXREG (Hedeker and Gibbons, 1996).

3.3. Parameters estimates

3.3.1. Estimation without covariates

First let us consider the models with a linear trend only; the estimated trend is plotted in Figures 3.1 and 3.2 . As expected, a decreasing behavior with a more rapid slope for the FS indicator may be observed. This result confirms the trend observed in Table 3.3 and can be explained by two factors: first of all, a higher improvement of household conditions in respect of the economic condition in the reference period; secondly, the FM indicator decrease across Waves 1-6 and a slight increase in Wave 7, making the slope smoother. The slope pendencies vary across macro regions: this confirms the presence of geographical differences. Parameter estimates for the residual structure are reported in Table 3.4 for the whole sample[7] and the three macro-regions[8].

For the moment, we are interested in comparisons among regions; comparison differences between measures will be considered in section 3.4. Let us consider the variation components in the FM poverty indicator for the whole sample. The total within year variance is about 0.98.

[6] Covariate sample means calculated for repeated observations of individuals aged 16+ in each of the three macro-regions are reported in the Appendix (Table .5).

[7] $\widehat{\gamma}$ is the residual interclass correlation coefficient computed by $\widehat{\sigma}_{\delta}^2/\widehat{\sigma}_u^2$. It can be interpreted as the fraction of total variability that is due to individual level.

[8] The estimates of the trend parameters and variance components with relative standard errors are reported in the Appendix (Tables .2-.4). It is important to underline that the use of this standard error to perfom hypothesis tests for the variance components is problematic (Bryk and Raudenbush, 1992), thus the significance of this component has been tested by the likelihood ratio test.

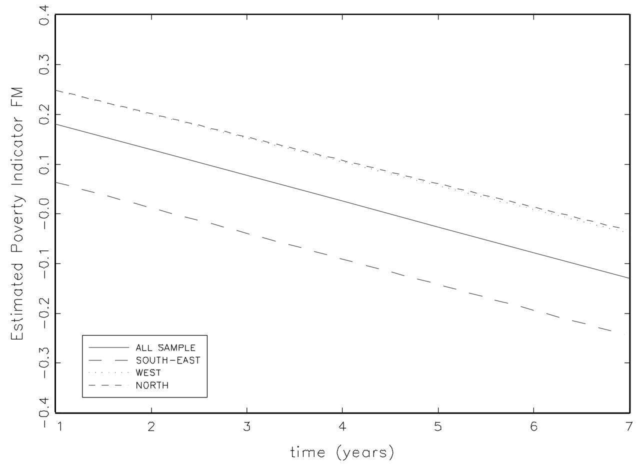

Figure 3.1: Estimated trend across macro-regions (FM measure)

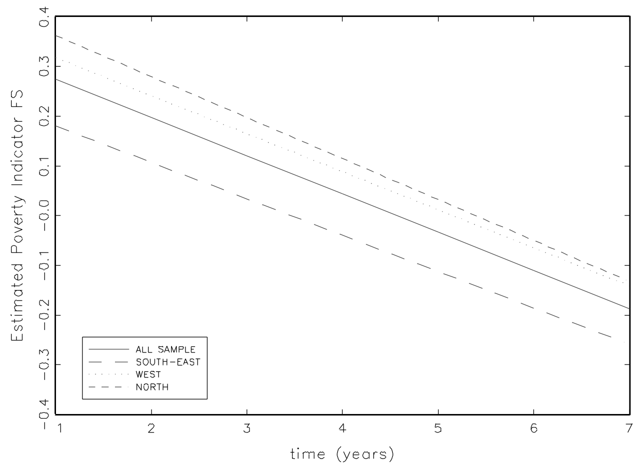

Figure 3.2: Estimated trend across macro-regions (FS measure)

Table 3.4: Components of variance; autocorrelated individual component models

		$\widehat{\sigma}^2_u$	$\widehat{\sigma}^2_\delta$	$\widehat{\sigma}^2_\xi$	$\widehat{\sigma}^2_\eta$	$\widehat{\gamma}$	$\widehat{\rho}$
FM	All	0.980	0.570	0.410	0.346	0.622	0.348
	South-East	1.109	0.656	0.454	0.395	0.624	0.360
	West	0.837	0.490	0.347	0.312	0.611	0.320
	North	0.853	0.494	0.359	0.313	0.613	0.358
FS	All	0.959	0.683	0.276	0.193	0.780	0.548
	South-East	0.994	0.707	0.287	0.205	0.775	0.537
	West	0.921	0.681	0.240	0.177	0.513	0.793
	North	0.924	0.629	0.295	0.190	0.768	0.598

Permanent poverty indicator differences among individuals represent 62% per cent of total variation. Of the 38% per cent remaining stochastic variation from period to period, 35% percentage points may be considered purely stochastic variation, the remainder being due to serial correlation. The results show a larger permanent and transitory variance in poverty indicators among people living in the South-East in respect of the other two regions.

According to the FS poverty indicator the total variance is explained for 78% by the permanent component and for about 20% by purely stochastic variation. Permanent and transitory variations seem to be different in Western regions with respect to South-Eastern and Northern regions.

In conclusion, results show differences in the variation components, among regions. This fact suggests maintaining the geographical partition in order to estimate models with observable covariates.

3.3.2. Estimation with covariates

Maximum likelihood estimates of the parameters for each region are reported in the Appendix (Tables .5-.7). In each model the dependent variable is the poverty indicator $y_{it}^{(\bullet)}$: thus a positive sign for the parameter of a significant covariate, corresponds to a higher deprivation risk.

Consider the role of the measured variables (fixed effects) in the variation of the membership function.

Observing Tables .5-.7 in the Appendix we note that for a subset of the covariates considered in the analysis the effect remains unchanged across the regions and there are no differences between the measures FS and FM. The variables belonging to the mentioned subset are: AGE and AGE2, JBSTA1, QUAL1 and QUAL2, MASTA, CH0_4. For these variables, the effect is the one expected. The household head age has a quadratic effect on the degree of deprivation, with a minimum at the age of about fifty years (for the FM measure this is coherent with the life-cycle theory). The poverty indicator is lower if the head of the household is employed or self-employed and the degree of deprivation tends to decrease as the educational level increases. Heads of household married or in couples make the membership function smaller than other marital status; such an effect is likely to be associated with the age of the head of household and/or with more than one earner in the household; the presence of children aged 0-4 years is a deprivation risk.

Although the effect of children and of household size will be analyzed in more depth in section 3.4, we add some other comments regarding the effect of the variable sex of the head of household and the differences between the three macro-regions.

The SEX variable is not always significantly different from zero; however, when its effect is significative, it is negative; that is, households headed by men are advantageous. Particularly with respect to the FM measure, the SEX variable has no effect in the North and West regions, while it has the usual negative effect in the South-East. Analysing poverty with the FS measure, the effects of the children over four years across the regions are also different: the presence of children aged 5-11 has no effect in the regions of the South-East and the West, while it has positive effect in the Northern area, making the membership function higher; vice versa, the presence of children 12-18 has no effect in the Northern area and has a negative effect in the other regions. An explanation could be that the Northern region seems to be poorer than the others, according to the FM measure (this trend is more evident in the comparison between the North and the South East); therefore it is likely that, children aged 12-18 living in the Northern Region, ask

Table 3.5: Components of Variance, Autocorrelated Individual Component Models

		$\widehat{\sigma}_u^2$	$\widehat{\sigma}_\delta^2$	$\widehat{\sigma}_\xi^2$	$\widehat{\sigma}_\eta^2$	$\widehat{\gamma}$	$\widehat{\rho}$
	SOUTH-EAST	0.803	0.392	0.411	0.366	0.517	0.329
FM	WEST	0.601	0.300	0.310	0.286	0.512	0.285
	NORTH	0.569	0.251	0.318	0.287	0.466	0.314
	SOUTH-EAST	0.602	0.378	0.224	0.166	0.695	0.508
FS	WEST	0.580	0.383	0.197	0.154	0.713	0.467
	NORTH	0.532	0.318	0.214	0.155	0.672	0.526

for less than their contemporaries living in the other regions.

As has been done for the models with trend, in Table 3.5 components of variance are reported. Using the likelihood ratio test, all these parameters are significantly different from zero. Generally, the variances $\widehat{\sigma}_u^2$ are smaller than those in Table 3.4; in fact, the observable variables inserted in the models explain part of the total variation. The permanent component explains a large part of the total error variance, both in the FM and the FS measure and in each macro-region . This result suggests that whatever measure is used, the effect of unobserved heterogeneity, interpreted as the effect of permanent differences among individuals, plays an important role in the analysis of poverty dynamics in each macro-region. Autocorrelation has an evident effect as well, it shows the smallest effect in Western regions according to both measures; on the contrary the permanent error component is the highest in this macro-region.

The correlation of poverty indicators for both measures in adjacent years across individuals $(\widehat{\gamma} + (1 - \widehat{\gamma})\,\widehat{\rho})$ declines as the time distance increases in each macro-region. Such a decrease is lower in the Western region and is generally stronger for the FS measure.

3.4. Comparison between measures

Remembering that the main objective of the paper lies in comparing the two alternative measures for the definition of the concept of poverty; we aim at showing that the meausures are, in effect, complementary to each other.

Let us consider again the effect of the sex of the head of household: it seems to be more effective on the measurement of non monetary poverty. Thus the deprivation for households headed by a woman, is likely to be more accentuated in terms of lack of durable goods or amenities in the household than in terms of income.

In the FM measure, the effect of the variable JBSTA2 is, as expected, always positive; the effect of an unemployed head of household is not so obvious according to the FS measure: in the South- East and West such an effect is not significantly different from zero, while in the North the effect is negative. It seems that according to the FS, unemployment does not affect material deprivation. The different effect of JBSTA2 regarding the FM and FS measure is likely to be related to the volatility of the income with respect to durable goods or housing conditions.

It has been previously specified that the presence of children, independently of the age, makes the household poorer according to the FM meausure; the effect of children is more complex considering the FS meausure. The presence of children aged 0-4 years has a positive effect (as in the FM measure); this effect could be associated with the generally young age of the head of household and with the fact that children aged 0-4 years are too young to demand durable goods. The presence of children aged 5-11 is not significant in two of the three regions; on the contrary, the effect of the presence of children aged 12-18 is very important. In two of the three regions the effect is negative, that is, it corresponds to lower deprivation. Such an effect could also associated with the following factors: the age of the parents: the fact that children aged 12-18 could "impose" on their parents to buy some of the durable goods considered in the analysis: the head of household is in the age range of minimum monetary poverty and therefore has the possibility to buy a house.

A quadratic specification of the household size is significant according to the FS measure; the membership function decreases with the increase of the household size up to five members. Beyond five members, it seems that there are not sufficient economic resources to face the needs of the household members.

Monetary deprivation generally increases as the household size increases. In the South East regions, only the linear specification of SIZE is significant with a positive effect; the quadratic specification is significant in the other regions with a minimum of 6.2 members in the North and 7.6 members in the West. It is reasonable to think that the increasing trend of the membership function up to 6-7 household members is associated with the increasing number of children. However it is quite curious that beyond 6-7 members the membership function tends to decrease; a possible explanation could be the presence of one or more adult members in the household contributing in monetary terms.

Now let us compare the variance components in the two different measure used, reported in Table 3.5. The main difference consists in the autoregressive components: according to the FS measures the autocorrelation coefficients are larger than those in the FM. Although a serial correlation model with time-dependence variables cannot be interpreted as a state-dependence model (Lindsey, 1999), it is plausible that the residuals at time $t-1$ have a higher impact on actual poverty in the FS measure. This is because housing conditions and possession of durable goods are much less volatile than monetary variables.

Permanent component patterns (among regions) are similar in the two measures; this is plausible since permanent components capture the effect of permanent differences among individuals.

Differences in the residual error component are more evident; this component includes either the effect of transitory variables or measurement errors. The two components cannot be distinguished here; anyway, the higher residual error component in the FM meausure can be explained by a larger incidence of measurement errors.

4. Some final remarks

In conclusion, the two measures lead to quite different results; the high residual variability in the FM measure suggests that the FS measure can be used to complement the picture: the simultaneous use of the two measures can help to comprehend the phenomenon of deprivation better.

The most important differences between the two measures in the empirical analysis regard: the slope of trend, the effect of household size and the presence of children aged 4+; the effect of autocorrelation.

From a methodological point of view, it can be added that the model specified can be generalised; for instance some of the hypotheses made, such as stationarity and independence between unobserved heterogeneity and covariates, can be relaxed.

In our analysis, we assume that labour force status affects the poverty condition, since this assumption could be considered quite strong. Another interesting issue to be analysed relates to the interdependence between the individual labour force process and the poverty process.

It is also important to remark that it would be interesting to consider a dummy for household changes such as covariate in the models, since these changes could influence the poverty process.

Finally, we have to reflect on the longitudinal unit of analysis. Assigning the same poverty indicator to each member of the household, as has been done, is a rather strong hypothesis, since it implies that household members share the same standard of living. Anyway, such a hypothesis seems to be more reasonable than considering each household member independent from the others. In fact, according to the latter measure, an individual poverty indicator should be calculated for each member of the household without considering the strong relation existing between household members. The problem connected with this assumption is that observations on individuals within households are not independent and so traditional standard error estimates are biased.

It is also important to underline another methodological issue: the model specified in Section 2 takes into account an individual random effect even if the dependent variable has been defined on the basis of the household indicator; however, the specification of a household random effect would lead again to the problem of the definition of the longitudinal unit of analysis, introduced above.

For this reason, it could be interesting to examine this problem closely. In fact, in order to correct the standard error estimates the balanced half-sampled technique (Wolter, 1985) should be used; otherwise multilevel models could be specified for dealing with nested individual observations in households (Snijders and Bosker, 1999; Goldstein, 1995).

References

ANDERSON T.W., HSIAO C. (1981), Estimation of dynamic models with error components, *Journal of the American Statistical Association*, 76, pp. 598-606.

ANDERSON T.W., HSIAO C. (1982), Formulation and estimation of dynamic models using panel data, *Journal of Econometrics*, 18, pp. 47-82.

BANE M.J., ELLWOOD D.T., (1986), Slipping into and out of poverty: the dynamics of spells, *Journal of Human Resources*, 21, pp. 1-23.

BARDASI E., JENKINS S.P., RIGG J.A. (1999), Documentation for Derived Current and Annual Net Household Income Variables, BHPS waves 1-7, *Working Papers of the ESRC Research Centre on Micro-Social Change.* Paper 99-25, Colchester: University of Essex.

BETTI G., VERMA V. (2000), Measuring the degree of poverty in a dynamic and comparative context, for presentation to the 5th International Meeting on Quantitative Methods in Applied Sciences, Siena, May 31 - June 3, 2000.

BRYK A. S., RUDENBUSH S.W. (1992), *Hierarchical linear models: Applications and data analysis methods*, Newbury Park, CA: Sage Publications, Inc.

BUCK N., ERMISCH J.F., JENKINS S.P. (1996), Choosing a longitudinal survey design: the issues, *Occasional Papers of the ESRC Research Centre on Micro-Social Change.* Paper 96-1, Colchester: University of Essex.

CERIOLI A., ZANI S. (1990), A fuzzy approach to the measurement of poverty, in Dagum C., Zenga M. (eds.), Income and Wealth Distribution, Inequality and Poverty, (proc. Pavia, Italy), *Studies in Contemporary Economics*, Springer Verlag, Berlin, pp.272-284.

CHELI B. (1995), Totally fuzzy and relative measures in dynamics context, *Metron* 53, pp.183-205.

Cheli B., Lemmi A. (1995), A totally fuzzy and relative approach to the multidimensional analysis of poverty, *Economic Notes*, 24, pp. 115-134.

Dagum C. (1977), A new model for personal income distribution: specification and estimation, *Economie Appliquée*, 3, pp. 413-436.

Dubois D., Prade H. (1980), *Fuzzy sets and systems*, Academic Press, Boston, New York, London.

Duncan G.J. (1983), The Implications of changing family composition for the dynamic analysis of family economic well-being, in Atkinson A.B., Cowell F.A. (eds.), Panel Data on Incomes, *Occasional Paper, 2, STICERD*, London School of Economics, London.

Duncan G.J., Coe R.D., Hill M.S. (1984), The dynamics of poverty, in Duncan D.J., Coe R.D., Corcoran M.E., Hill M.S., Hoffman S.D., Morgan J.N. (eds.), *Years of Poverty, Years of Plenty: the Changing Economic Fortunes of American Workers and Families*, Institute for Social Research, University of Michigan, Ann Arbor.

Goldstein H., Healy M.J.R., Rasbash J. (1994), Multilevel time series models with applications to repeated measures data, *Statistics in Medicine,* 13, pp. 1643-1655.

Goldstein H. (1995), *Multilevel statistical models*, New York: Halstead Press.

Hedeker D., Gibbons R.D. (1996), MIXREG: a computer program for mixed-effects regression analysis with autocorrelated errors, *Computer Methods and Program in Biomedicine.*

Jenkins S.P. (1999), Modelling household income dynamics, *Working Papers of the*

ESRC Research Centre on Micro-Social Change. Paper 99-1, Colchester: University of Essex.

JONES R.H. (1993), *Longitudinal data with serial correlation: a state-space approach*, Chapman & Hall.

LINDSEY J.K (1999), *Models for repeated measurements*, Oxford University Press.

LILLARD L.A., WILLIS R.J. (1978), Dynamic Aspect of Earning Mobility, *Econometrica*, 46, pp. 985-1011.

MANSOUR H., NORHEIM E.V., RUTLEDGE J.J. (1985), Maximum likelihood estimation of variance components in repeated measure designs assuming autoregressive errors, *Econometrics*, 41, pp. 287-294.

RAGHUNATHAN T.E., LEPKOWSKI J., VAN VOEWYK J., SOLENBERGER P. (1997), *A multivariate techinique for imputing missing values using a sequence of regression models*, Technical Report, Survey Methodology Program, Survey Research Center ISR, University of Michigan.

SNIJDERS A.B., BOSKER R.J. (1999), *Multilevel analysis: an introduction to basic and advanced multilevel modeling*, Sage Publications: London.

STEVENS A.H. (1995), Climbing out of poverty, falling back in: measuring the persistence of poverty over multiple spells, *NBER Working Paper Series*, Working Paper 5390, National Bureau of Economic Research, Cambridge, MA.

TAYLOR A. (1994), Appendix: sample characyeristics, attrition and weighting, in Buck N., Gershuny J., Rose D., Scott J. (eds.), *Changing household: the British Household*

Panel Survey 1990-1992, Colchester: University of Essex.

TAYLOR M.F., BRICE J., BUCK N., PRENTICE E. (1996), *British Household Panel Survey user manual volume A: introduction, technical report and appendices*, Colchester: University of Essex.

TAYLOR M.F. (1998), *British Household Panel Survey user manual: introduction, technical report and appendices*, Colchester: University of Essex.

TRIVELLATO U. (1998), Il monitoraggio della povertà e della sua dinamica: questioni di misura e evidenze empiriche, *Statistica*, 58, pp. 549-575.

WOLTER K.M. (1985), *Introduction to variance estimation*, Springer Series in Statistics.

ZADEH L.A. (1965), Fuzzy Sets, *Information and control*, 8, pp.338-353.

Table .1: Mean, (standard deviation) of variables

	REGIONS			
	South-East	**West**	**North**	**Total**
SEX	0.782	0.761	0.733	0.762
AGE	47.707 (16.44)	49.068 (16.71)	47.332 (17.14)	47.992 (16.75)
JBSTA1	0.658	0.614	0.611	0.631
JBSTA2	0.053	0.047	0.047	0.496
QUAL1	0.189	0.153	0.176	0.175
QUAL2	0.438	0.433	0.417	0.430
MASTA	0.727	0.735	0.712	0.725
CH0_4	0.120	0.123	0.133	0.125
CH5_11	0.174	0.189	0.171	0.177
CH12_18	0.199	0.175	0.167	0.182
SIZE	2.931 (1.36)	2.810 (1.34)	2.757 (1.32)	2.803 (1.34)
# individuals	5322	3657	3787	12766
#repeated observations	23809	16891	17264	57964

Table .2: Marginal Maximum Likelihood estimates, model with linear trend: South-Eastern Region

	FM measure		**FS measure**	
Variables	Estimates (s.e)		Estimates (s.e)	
Fixed effects				
INTERCEPT	0.1146 (0.016)	***	0.2527 (0.015)	***
PEPI	-0.0515 (0.002)	***	-0.0731 (0.002)	***
$\widehat{\rho}$	0.3601 (0.010)	***	0.5377 (0.010)	***
Random effect				
$\widehat{\sigma}_{\delta}^2$	0.6569 (0.017)	***	0.7079 (0.017)	***
$\widehat{\sigma}_{\eta}^2$	0.3958 (0.004)	***	0.2059 (0.002)	***
logL	-27174.141		-20629.992	

Table .3: Marginal Maximum Likelihood estimates, model with linear trend: Western Region

	FM measure		**FS measure**	
Variables	Estimates (s.e)		Estimates (s.e)	
Fixed effects				
INTERCEPT	0.3329 (0.017)	***	0.3938 (0.017)	***
PEPI	-0.0556 (0.002)	***	-0.0764 (0.002)	***
$\widehat{\rho}$	0.3209 (0.011)	***	0.5139 (0.012)	***
Random effect				
$\widehat{\sigma}_{\delta}^2$	0.4902 (0.015)	***	0.6810 (0.019)	***
$\widehat{\sigma}_{\eta}^2$	0.3124 (0.003)	***	0.1779 (0.002)	***
logL	-17205.701		-13524.282	

Table .4: Marginal Maximum Likelihood estimates, model with linear trend: Northern Region

	FM measure		FS measure	
Variables	Estimates (s.e)		Estimates (s.e)	
Fixed effects				
INTERCEPT	0.2955 (0.017)	***	0.44385 (0.017)	***
PEPI	-0.0478 (0.003)	***	-0.0824 (0.003)	***
$\widehat{\rho}$	0.3582 (0.011)	***	0.5982 (0.012)	***
Random effect				
$\widehat{\sigma}^2_\delta$	0.4948 (0.015)	***	0.6293 (0.019)	***
$\widehat{\sigma}^2_\eta$	0.3130 (0.003)	***	0.1905 (0.002)	***
logL	-17585.688		-14115.052	

Table .5: Marginal Maximum Likelihood estimates: South-Eastern Region

	FM measure		FS measure	
Variables	Estimates (s.e)		Estimates (s.e)	
Fixed effects				
INTERCEPT	1.4900 (0.069)	***	3.7425 (0.054)	***
PEPI	-0.0469 (0.002)	***	-0.0811 (0.002)	***
SEX	-0.0630 (0.020)	***	-0.0164 (0.015)	
AGE	-0.0417 (0.003)	***	-0.0957 (0.002)	***
AGE2	0.0004 (0.000)	***	0.0009 (0.000)	***
JBSTA1	-0.4498 (0.187)	***	-0.1362 (0.013)	***
JBSTA2	0.2857 (0.026)	***	0.0214 (0.017)	
QUAL1	-0.7494 (0.027)	***	-0.4548 (0.022)	***
QUAL2	-0.3043 (0.022)	***	-0.2981 (0.018)	***
MASTA	-0.1908 (0.021)	***	-0.2773 (0.015)	***
CH04	0.2171 (0.021)	***	0.0872 (0.015)	***
CH511	0.2223 (0.020)	***	-0.0076 (0.015)	
CH1218	0.1698 (0.015)	***	-0.0434 (0.010)	***
SIZE	0.0722 (0.022)	***	-0.4145 (0.016)	***
SIZE2	-0.0003 (0.003)		0.0410 (0.002)	***
$\widehat{\rho}$	0.3295 (0.010)	***	0.5086 (0.010)	***
Random effect				
$\widehat{\sigma}^2_\delta$	0.3925 (0.011)	***	0.3789 (0.010)	***
$\widehat{\sigma}^2_\eta$	0.3667 (0.003)	***	0.1660 (0.001)	***
logL	-25466.619		-17246.878	

Table .6: Marginal Maximum Likelihood estimates: Western Region

	FM measure		**FS measure**	
Variables	Estimates (s.e)		Estimates (s.e)	
Fixed effects				
INTERCEPT	1.2442 (0.076)	***	3.2812 (0.065)	***
PEPI	-0.0547 (0.003)	***	-0.0863 (0.002)	***
SEX	-0.0289 (0.021)		-0.0991 (0.017)	***
AGE	-0.0320 (0.003)	***	-0.0764 (0.003)	***
AGE2	0.0003 (0.000)	***	0.0008 (0.000)	***
JBSTA1	-0.4693 (0.019)	***	-0.1122 (0.015)	***
JBSTA2	0.2312 (0.027)	***	0.0370 (0.020)	*
QUAL1	-0.5818 (0.029)	***	-0.4465 (0.027)	***
QUAL2	-0.1774 (0.022)	***	-0.2293 (0.020)	***
MASTA	-0.2531 (0.024)	***	-0.2462 (0.019)	***
CH04	0.1422 (0.022)	***	0.0477 (0.017)	***
CH511	0.1978 (0.021)	***	-0.0038 (0.017)	
CH1218	0.1335 (0.016)	***	-0.0297 (0.017)	**
SIZE	0.1531 (0.027)	***	-0.3786 (0.023)	***
SIZE2	-0.0102 (0.004)	***	0.0376 (0.003)	
$\widehat{\rho}$	0.2852 (0.011)	***	0.4677 (0.012)	***
Random effect				
$\widehat{\sigma}_{\delta}^{2}$	0.3001 (0.010)	***	0.3838 (0.012)	***
$\widehat{\sigma}_{\eta}^{2}$	0.2866 (0.003)	***	0.1545 (0.001)	***
logL	-15979.503		-11739.099	

Table .7: Marginal Maximum Likelihood estimates: Northern Region

	FM measure		**FS measure**	
Variables	Estimates (s.e)		Estimates (s.e)	
Fixed effects				
INTERCEPT	1.4908 (0.071)	***	3.75923 (0.062)	***
PEPI	-0.0449 (0.003)	***	-0.0892 (0.002)	***
SEX	-0.0040 (0.020)		-0.0878 (0.016)	***
AGE	-0.0388 (0.003)	***	-0.0956 (0.003)	***
AGE2	0.0004 (0.000)	***	0.0009 (0.000)	***
JBSTA1	-0.5412 (0.018)	***	-0.1536 (0.015)	***
JBSTA2	0.1833 (0.027)	***	-0.0541 (0.020)	***
QUAL1	-0.6270 (0.027)	***	-0.3579 (0.025)	***
QUAL2	-0.1790 (0.021)	***	-0.2273 (0.019)	***
MASTA	-0.2786 (0.022)	***	-0.3069 (0.018)	***
CH04	0.1393 (0.021)	***	0.0532 (0.017)	***
CH511	0.2220 (0.020)	***	0.0342 (0.016)	**
CH1218	0.0935 (0.016)	***	0.0199 (0.012)	
SIZE	0.1460 (0.026)	***	-0.3551 (0.021)	***
SIZE2	-0.0117 (0.004)	***	0.0327 (0.003)	***
$\widehat{\rho}$	0.3142 (0.011)	***	0.5267 (0.011)	***
Random effect				
$\widehat{\sigma}_{\delta}^{2}$	0.3186 (0.010)	***	0.2515 (0.009)	***
$\widehat{\sigma}_{\eta}^{2}$	0.2878 (0.003)	***	0.1554 (0.001)	***
logL	-16115.297		-11731.096	