IRISS at CEPS/INSTEAD

An Integrated Research Infrastructure in the Socio-Economic Sciences

A LATENT CLASS APPLICATION TO THE MEASUREMENT OF POVERTY

by

Pasi Moisio

IRISS WORKING PAPER SERIES

No. 2001-08

Please pay a visit to http://www.ceps.lu/iriss/iriss.htm

IRISS-C/I

An Integrated Research Infrastructure in the Socio-Economic Sciences at CEPS/Instead, Luxembourg

A project supported by the European Commission, the Ministry of Culture, Higher Education and Research (Luxembourg) and the National Science Foundation (USA)

RESEARCH GRANTS

for individual or collaborative research projects
(grants awarded for periods of 4-6 weeks)

Who may apply?

IRISS-C/I supports collaborative and/or internationally comparative research in economics and other social sciences. We encourage applications from all interested individuals who want to carry out research in the fields of expertise of CEPS/Instead. A separate funding agreement is in place for US researchers. Their access to IRISS-C/I is funded by the National Science Foundation.

What is offered by IRISS-C/I?

Free access to the IRISS-C/I research infrastructure (office, computer, library...);
Access to the CEPS/Instead archive of micro-data;
Technical and scientific assistance;
Free accommodation and a contribution towards travel and subsistence costs.

Research areas

Survey and panel data methodology; income and poverty dynamics; persistent and new poverty; gender, ethnic and social inequality; unemployment; segmentation of labour markets; new forms of work/a-typical employment; education and training; social protection and redistributive policies; impact of ageing populations; intergenerational relations; effects of family-related policies; regional development and structural change.

Additional information and application forms

IRISS-C/I
B.P. 48 L-4501 Differdange (Luxembourg)
IRISS-C/I homepage: http://www.ceps.lu/iriss/iriss.htm

IRISS-C/I Working Paper:

A LATENT CLASS APPLICATION TO THE MEASUREMENT OF POVERTY[1]

by
Pasi Moisio
European University Institute, Florence
pasi.moisio@iue.it

[1] The author wishes to thank IRISS-C/I at CEPS/INSTEAD (Luxembourg) for its financial and academic support.

INTRODUCTION

Poverty is an object of normative, administrative and methodological interest for various actors. This means that it is inevitably a political concept and therefore, *per se*, continuously debated (Alcock 1993, 3). In very general terms poverty can be defined as living at the bottom of welfare distribution. Hence, poverty is closely related to inequality, though it is not the same as inequality: if a society is equal but the standards of living are low across the society, it follows that everyone in the society is poor, not that there is no poverty. The reason why special attention is paid to the bottom of welfare distribution is the belief that there is a level of welfare below which people suffer some form(s) of deprivation (see Creedy 1998, 25). This is why the concept of poverty threshold (or poverty line, cut-off point, etc.) has a central place in definitions of poverty and in poverty research in general. The poverty threshold represents the fundamental idea of the poverty measurement that there is a threshold in welfare distribution below which well-being drops sharply - so much so that it is reasonable to think of it as a qualitative difference.

There are numerous poverty measurements in the poverty literature which identify poor persons or households by low resources, low expenditures, poor living conditions or subjective feelings. The main problem, however, is how to select what indicators should be used in different countries i.e. how to judge what are 'good' poverty indicators and what are not, because different poverty indicators meet only one assumption of nominally parallel measurements. In other words, they are operationalisations of the same concept and using the internal consistency of indicators, for example, as a criterion for item selection does not work. Thus, we need an item selection method for 'quasi' parallel poverty measurements, one that can handle indicators that are not nominally parallel.[2] The logic behind using a set of parallel measurements is that this way the error of measurement can be estimated, which is not possible if only one indicator is used.

In this paper, I suggest that the unrestricted latent class model (LCM) and the axiom of local independence can offer a method for selecting and evaluating different poverty indicators and thus make possible multidimensional poverty measurement, where poverty is measured using several indicators side-by-side. In other words, the LCM can be used to select the best possible set of poverty indicators that are, to some extent, comparable across different countries. In this way, for the first time we do not have to rely only on the theory and substance when constructing or selecting poverty indicators - with the LCM we can also test whether different operationalisations of poverty really measure the same latent phenomenon. In addition, two other

[2] A detailed explanation on (nominally) parallel measurements can be found in Bollen (1989). For my purposes here it is enough to know that nominally parallel measurements are the set of indicators that measure same concept and share certain statistical characteristics, namely, that the error of measurement is equal and uncorrelated across measurements.

ways of applying the LCM in the measurement of poverty are presented briefly in the paper: how to construct a sample and indicator-free latent poverty index and how to use the ordinal LCM as an alternative for the summary index of deprivation.

The paper starts with the presentation of development in the measurement of poverty and deprivation. The purpose of this review is to show how the application of latent class analysis is a natural continuation to the latest trend in the measurement of poverty, where poverty is understood and measured as a multidimensional phenomenon.

DEVELOPMENT IN THE MEASUREMENT OF POVERTY AND DEPRIVATION

Three historical stages in poverty measurement can be distinguished (Ringen 1985). However, these three approaches all have their own supporters and they co-exist in current poverty research, so to talk about three clearcut stages may be slightly exaggerated.

These approaches and their development can be briefly described as follows. The modern measurement of poverty can be seen as starting in Rowntree's pioneering work in York at the beginning of the 20th century, which laid the first conceptual and theoretical foundation for the measurement of poverty. His definition of poverty as minimum subsistence and the food-basket method for measuring it dominated the measurement of poverty for almost a century. It was not until the beginning of the 1970's that this *absolute poverty approach* was replaced by a relative definition of poverty. Peter Townsend (1979,1987) was the leading name in this new *relative poverty approach* and he defined it as a lack of material resources to maintain an adequate way of living in a given society. Although the definition of poverty changed from absolute to relative, i.e. from a biological necessity as the poverty threshold was replaced by a threshold calculated from the distribution of material resources, the measurement of poverty itself remained the same - it was measured as a lack of material resources.

However, there was a logical break between the definition of poverty and indicators that were used for measuring it. In other words, poverty was measured as the lack of material resources, but the definition of poverty referred to standards of living. This contradiction gave rise to criticism in the 1980's, especially after several studies showed that the link between low material resources and low standards of living was much more complex than a simple causal one: it was shown that equal material resources do not necessarily result in equal welfare and that material resources are not necessarily a function of welfare. The leading critic against the relative poverty approach was Stein Ringen (1985), who believed that poverty should be measured directly as poor living conditions, i.e. deprivation, and not indirectly as the

lack of material resources. This led to the third and new *relative deprivation approach* of the measurement of poverty.[3]

Berghman (1995, 21) has offered another analytical framework for the measurement of poverty by clarifying relationships between a concept of poverty and deprivation (and also social exclusion[4]). According to him, poverty and deprivation are static outcomes and the latter differ from the former in that deprivation is multidimensional where poverty, i.e. lack of material resources, is one-dimension. Social exclusion, on the other hand, is a concept that refers to a multidimensional process, where impoverishment is one dimension. Thus, poverty is one dimension in deprivation and impoverishment is one dimension in social exclusion. However, there is a good deal of confusion in the literature of poverty research whether poverty should be seen as a lack of resources or poor living conditions (Nolan & Whelan 1996, 10-14). Measuring poverty as low resources is often referred to as the indirect measurement of poverty, and measuring poverty as poor living conditions as the direct measurement of poverty (Ringen 1985).

Thus, the relationship between deprivation and poverty can be seen in different ways, though they are not exclusive conceptualisations. Deprivation can be seen as the manifestation of poverty (Townsend and Ringen) or poverty can be seen as one dimension in deprivation (Berghman). In recent years, poverty has been more conceptualised as one dimension in deprivation. However, as mentioned above, these two conceptualisations of poverty can (and do) co-exist and there is no logical contradiction between them. Hence, we can assume that poverty is one dimension in deprivation that cannot be directly observed. This is why we are forced to measure poverty indirectly by measuring the other dimensions of deprivation, such as poor living conditions.

The second division along with the division between direct and indirect measurements is that between aggregated indicators and simple head count measurements (Sen 1979). Sociologists mainly use head count indicators, which identify and calculate the number of poor persons and families in society. Most of the poverty measurement literature among economists after 1970's has concentrated on developing more complex poverty indexes, often together with measurements of inequality like the Gini coefficient or Lorenz distribution (Creedy 1998, 13-17). However, it seems to be that the simple head-count poverty indicators have kept their position as the main type of poverty measurement. This is perhaps because it is usually adequate to answer research questions about poverty: locating poverty in society is the main information that, for example, social political decision-making and administration usually need. From their perspective, the crucial difference be-

[3] It was Ringen (1985) who originally presented these three stages in the measurement of poverty.

[4] One could say that the concept of social exclusion has superseded 'old' concepts of poverty and deprivation, at least in political rhetoric. Although these three terms are often used superficially as synonyms for each other, the concept of social exclusion carries some implicit assumptions about the nature of social risks that can have influence on social policy and public opinion (Moisio 2001).

tween different poverty measurements is the underlying definition of poverty behind the indicator, not the details or construction of the indicator (Ruggles 1990, 14).

These questions are not the only questions that have caused disputes in poverty research. Questions about the poverty threshold and the unit of analysis also cause constant debate. In the cut-off points that divide a population into poor and non-poor, fixed and relative thresholds are often separated. A fixed poverty line is set before measurement, using some additional information outside the sample, experts, administration, etc. deciding on this fixed poverty line. The most influential fixed poverty line is probably 'being a recipient of social assistance' and the U.S. Census' official poverty line (see Citro & Michael 1995, 24). The relative poverty line is defined as the function of some distribution. In most cases, this poverty line is the function of the income distribution in the sample: probably the most widely used is the relative income poverty line, as the net income less than 50% (or 60%) of the median of all households (see e.g. Eurostat 1998).

In sociological poverty research, poverty is usually measured at the family level and when individual level information is needed, persons in the same family are given the same value. Hence, the allocation of material welfare inside the household is (implicitly) assumed to base on reciprocal personal contacts and this way to be 'equal': this is not necessarily a valid assumption (Jenkins 1991). When using the household as the unit of analysis, we have to make comparable different-sized and composed household with different needs. A common method for doing this is to use an equivalisation scale to calculate the household's income or expenditure per consumption unit. The most widely used equivalisation scales for family incomes are the classical OECD scale and square root scale. Different equivalisation scales give somewhat different poverty figures (Atkinson 1995, 80-1.) Additionally, poverty rates at the household and family level are relatively unequal. Especially when comparing poverty figures between different countries with different family structures, results can be quite different whether we are using individual or family-level poverty rates.

After decades of disputes between different 'schools' on how poverty should be defined and measured, the trend in present-day sociological poverty research is towards a mixed methodology and multidimensional measurement. Poverty is defined and measured as a multidimensional phenomenon and different ways to measure poverty are seen as alternative ways to gain information on the same complex social problem. The multidimensional measurement of poverty is usually done with the set of indicators that include direct and indirect measurements as well as indicators measuring the subjective feeling of poverty. It is accepted that poverty is a multidimensional social problem that cannot be measured sufficiently or exhaustively with one indicator (see Kangas & Ritakallio 1998). The mixed methodology also means that the quantitative analysis of poverty is often supplemented with a qualitative analysis (or *vice versa*).

Nevertheless, there is one default in the multidimensional approach: in the current literature of poverty research, there is no test or method that ensures that selected poverty and deprivation indicators really measure the same latent phenomenon. This is a problem, since poverty and deprivation indicators are culturally biased - for example, the summary index about multiple problems in accommodation can be good a indicator of deprivation in one country, but not in another. Sometimes the internal consistency of the poverty and deprivation indicators is used to select indicators. But using an estimate like the Cronbach Alpha coefficient makes no sense when we are handling poverty and deprivation indicators that do not meet the assumptions of nominally parallel measurements: poverty indicators have different variances and the error of measurement is not random across indicators and persons. The most distinctive consequence of the fact that poverty indicators are not nominally parallel is that poverty indicators identify quite different parts of the population as poor (see Atkinson 1998).

We can, however, overcome anomalies caused by the non-parallelism of poverty indicators by constructing a latent structure model, where observed poverty indicators are handled as the imperfect manifestation of a latent poverty structure. Hence, different poverty and deprivation indicators do not meet the assumptions of parallel measurement, but they are operationalised to measure same concept: so we can say that the poverty indicators are 'quasi-parallel' measurements. Quasi-parallel poverty indicators reflect the latent structure from different angles and different sizes, but it is possible treat this since the model estimates parameters describing multiple sources of error.

The axiom of local independence provides a method for evaluating whether relationships between poverty and deprivation indicators in hand can be explained by some unmeasured latent structure. Because different poverty and deprivation indicators are supposed to measure the same thing - resources or living conditions so much below some threshold that we can speak about qualitative difference - they should be locally independent from each other in the latent classes of poor and non-poor. In other words, according to the rules of statistical inference, we should be able to construct a dichotomous latent variable behind the indicators of poverty and deprivation, and this model should meet the assumption of local independence. Otherwise, the multidimensional measurement of poverty is meaningless, since we have justified multidimensional measurement by stating that the actual measurement object is multidimensional poverty and we are just forced for practical reasons to measure only its different manifestations.

Thus I propose that the unrestricted latent class model should be used to test that different poverty and deprivation indicators really measure the same latent phenomenon - this way the selection of indicators is not relying

only on the theory and substance. However, there are also two other ways to apply the LCM in the measurement of poverty. First, the latent poverty classes can be used as an attempt to construct a sample and indicator-free poverty measurement. Gailly and Hausman (1980) have studied the possibility of constructing this kind of latent poverty variable using the latent trait model, namely, the Rasch model. Second, the ordinal LCM can be used to construct a latent ordinal deprivation index, instead of to sum up the scores of items for a summary index. Using the latent index one can treat the problem that different items measure deprivation with different sizes (and with different errors), one that Desai and Shah (1988) solved by weighting the items. However, presenting how the LCM can be used in these two other ways in the measurement of poverty requires another paper. In this paper only the first way, i.e. how the LCM can be used to select the best set of poverty and deprivation indicators, will be presented in detail.

LATENT CLASS MODEL AND LOCAL INDEPENDENCE

In general, latent structure models are defined as measurement models relating the discrete or continuous latent variable to the discrete scores or categories of manifest variables (Lazarsfeld and Henry 1968, 15-7). There can be more than one latent variable, however, though in the latent class model (LCM) and the latent trait model (LTM) usually only one latent variable is assumed. These two best-known latent structure models differ from each other in the respect that former construct latent categorical variable(s) behind the manifest variables, when latter construct continuous one(s). Relations between the latent variable and manifest variables are stochastic. In other words, latent structure models are statistical models, which means that relationships between latent and manifest variables are accounted for by probabilistic relationships, thus also allowing for the estimation of error in the model. These probabilistic relationships are treated under the axiom of local independence.

The axiom of local independence, formulated by Lazarsfeld and Henry (1968, 17), can be seen as the defining characteristics of latent structure models. In every latent structure model it is assumed that observed associations between manifest variables depend on the relationship between latent and manifest variables. Thus, local independence assumes that if we hold the latent variable constant, manifest variables should be statistically independent from each other (Heinen 1996, 6).

Presentation and estimation of the LCM is straightforward. Observed frequencies are reproduced by conditional probabilities using the following equation and the line of inference and model building is very similar to the building of a log-linear model.

E1: $$P(X = t \mid ABC|) = \pi_{ijkt}^{ABC\bar{X}} = \frac{\pi_{ijkt}^{ABCX}}{\pi_{ijk}^{ABC}}$$

The equation (E1) is the mathematical formalisation of the axiom of local independence for three manifest variables (A, B and C) and one latent variable (X). It equalises the conditional probabilities that a given case is located in a certain *ijk* cell in the observed three variable crosstabulation ABC when the latent class *t* of the latent variable X is given. Equation E1 represents a LCM with a one-dimensional latent distribution, but it can be developed for more than one latent variable.

Often the latent class model is presented in its more familiar log-linear form, shown in the equation E2. When the LCM presented in this way, it shows much better how close the LCM and log-linear model are to each other. The log-linear equation of LCM with three manifest variables (A, B and C) and one latent variable (X) is

E2: $$\log f_{ijkt}^{ABCX} = \theta + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX}$$

All the associations between manifest variables A, B and C are thus assumed to be conditionally independent when the latent variable X is introduced into the model.

The set of equations that yield maximum likelihood estimates for solving equation E1 (and equation E2 if using natural log transformation) is presented in Appendix A. For more detailed insights on the technical and philosophical foundations of the latent class model, see Goodman (1978), Lazarsfeld & Henry (1968) and McCutcheon (1987).

SELECTING POVERTY INDICATORS WITH THE LCM

To illustrate how the LCM and the axiom of local independence can be used as a tool to select poverty and deprivation indicators, I have done a small analysis with six poverty and deprivation indicators, using a data[5] from Finland, the Netherlands and the UK. The estimation of parameters was conducted using the LEM program (Vermunt 1997). The aim is to show how the LCM can be used to test whether selected six well-known poverty and deprivation indicators really measure the same phenomenon and how the information from this test can be used to select the best set of indicators for each country.

Thus, for a poverty or deprivation indicator to be selected it has to be, first (of course), justified theoretically and substantially, and secondly, it has to meet the assumption of local independence with other poverty indicators. In other words, local independence between different poverty and deprivation indicators is the requirement of statistical characteristics that these indicators have to pass, if we insist that they measure different dimensions or manifestations of one phenomenon - be it then called poverty or something else. Using culturally specific indicators like poverty and deprivation indicators, we can assure at least some kind of comparability across countries if the assumption of local independence is met. And even if the comparability is not completely assured, the indicators are easily interpreted, since no aggregations or weighting have taken place. In other words, once indicators have passed the test of local independence, they can be used as they are, side-by-side, to measure different manifestations of poverty. This is a simple, but transparent and easy to manage, method for measuring multidimensional poverty.

The six poverty and deprivation indicators are selected into the analysis according to their central position in the measurement of poverty. Their technical documentations are presented in Appendix B and their distributions in the three countries are shown in Table 1. Because variables A and F (recipient of social assistance or housing allowance and relative income poverty) are the two most important poverty measurements, the latter because of its publicity and wide use and the former because of its importance for social policy makers, they are required for the final set of indicators. Instead, the selection is made between variables B, C, D and E, where B and D can be seen as two indicators measuring the subjective feeling of poverty, and C and D indicators measuring poor living conditions. Thus, the LCM is used to study, whether it is better in some countries to measure the subjective feeling of poverty and the poor living conditions with different indicators.[6]

---

[5] Data is the third wave of European Commission Household Panel (ECHP) from the year 1996. The household surveys in ECHP are collected by national statistical offices or research institutes, and Eurostat is responsible for gathering and standardising them for comparative use (Eurostat 1999).

[6] Naturally, we could also create quasi-parallel indicators for relative income poverty, for example, and make a selection between them or include some other poverty indicators in

Table 1. Percentages of households identified as poor or deprived using different indicators

	Finland	Netherlands	The UK
A. Recipient of social assistance	11,9	8,2	20,4
B. Problems in paying rent or mortgage	11,8	2,0	6,8
C. Housing deprivation	9,7	13,6	18,7
D. Relative deprivation	5,7	4,3	7,0
E. Problems to make ends meet	17,9	14,6	18,4
F. Relative income poverty	8,0	8,3	11,3
N	4139	5175	3775

The unrestricted LCM with two latent classes was selected because it represents best the theoretical construction of the poverty threshold as a qualitative threshold dividing the population into two. In Table 2 the results from the selection of indicators in each country are presented. All of the models presented in the Table 2 failed to fit in the model with one latent class i.e. manifest variables are not statistically independent.

Table 2. Latent Class Models poverty and deprivation indicators

		Finland		Netherlands		The UK	
Model (T=2)	df	L-squared	P-value	L-squared	P-value	L-squared	P-value
A|X B|X C|X D|X E|X F|X	50	142,13	>0,001	125,91	>0,001	125,91	>0,001
A|X B|X C|X D|X F|X	20	44,47	0,001	38,33	0,008	65,96	>0,001
A|X B|X C|X E|X F|X	20	52,95	>0,001	53,79	>0,001	47,43	>0,001
A|X B|X D|X E|X F|X	20	69,84	>0,001	69,15	>0,001	50,46	>0,001
A|X C|X D|X E|X F|X	20	32,38	0,039	81,48	>0,001	57,23	>0,001
A|X B|X C|X F|X	6	12,83	0,046	5,05	0,538	24,52	>0,001
A|X B|X D|X F|X	6	20,47	0,002	16,47	0,011	8,22	0,222
A|X B|X E|X F|X	6	29,12	>0,001	21,52	0,002	11,72	0,068
A|X C|X D|X F|X	6	19,59	0,003	17,57	0,007	14,02	0,029
A|X C|X E|X F|X	6	5,68	0,460	24,23	>0,001	10,34	0,111
A|X D|X E|X F|X	6	11,21	0,082	37,66	>0,001	13,96	0,030

According the model fit diagnostics, it seems to be that the best set of indicators for Finland is A, C, E and F. This also means that indicators B and D do not 'work' in Finland. The explanation for why problems in paying scheduled rent or mortgage and the deprivation index calculated from the lack of some essential durables does not work in Finland has to be sought in our empirical knowledge. We also have to look for an explanation outside statistical mathematics for why the best set of indicators seems to be different in The Netherlands, where the best set according to the LCM is A, B, C and F. In the UK, there seem to be two possible combination of variables that met

---

the analysis. As stated above, this analysis is an exercise to show how local independence can be used to select indicators and this why the number of indicators is limited.

the axiom of local independence: A, B, D and F and the second set is same as in Finland; A, C, E and F.

Why is asking about problems in paying scheduled rents or mortgages and the durable deprivation index are good way of measuring the subjective feeling of poverty and the poor living conditions in The UK, but not in Finland? Why does it seem to be that multiple problems in accommodation are a good indicator to measure possible poor living conditions in Finland and in the Netherlands, but not in the UK? As I have said, these are questions where one has to seek answers outside mathematics, but the LCM is the only way of identifying this cultural biasness of some poverty and deprivation indicators. I do not propose that the axiom of local independence should be the only criteria when choosing indicators for measuring poverty multidimensionally - constructing indicators should always be guided by theory and common sense, but these indicators should meet the assumption of local independence as a guarantee that the indicators posses statistical characteristics of quasi parallel measurements. Without the assumption that the indicators measuring the different dimensions of poverty and deprivation are quasi-parallel, there are no grounds for insisting that we can measure poverty multidimensionally.

SUMMARY

In the paper, it has been proposed that latent class analysis can be used for testing that selected poverty and deprivation indicators really measure the same latent phenomenon and then using this information to select the best set of indicators. The axiom of local independence is presented as a requirement that different poverty and deprivation indicators (justified beforehand by the theory and substance) have to meet so that they cab be said to measure the same phenomenon - poverty. Using this method the best possible set of indicators was selected for Finland, the Netherlands and the UK.

According to the results, measuring the subjective feeling of poverty and poor living conditions should be carried out with slightly different questions in these three countries. Question 'is household able to make ends meet' seems to be a suitable indicator in Finland and in the UK, but not in the Netherlands. Also, measuring poor living conditions with the summary index of lack of some essential durables seems to work in the UK, but not in Finland or in the Netherlands, where a better indicator is the multiple problems in accommodation. Explanations for these results have to be sought outside statistical mathematics.

Before this test of local independence, constructing and selecting indicators for the measurement of poverty has relied on practically only theory and substance, since there have not been any statistical tests or methods for evaluating statistical characteristics of quasi-parallel poverty indicators. Using ordinal correlation is insufficient to study scalability and tests for in-

ternal consistency imposes statistical assumptions on indicators that poverty and deprivation indicators do not meet. Thus, I have proposed that we should use the unrestricted latent class model as a method to test if different poverty and deprivation indicators posses wanted statistical characteristics, namely, that they measure the same latent phenomenon. For this, all poverty and deprivation indicators have to be locally independent in the latent poor - non-poor classes.

Naturally, a statistical model cannot select indicators on behalf of the researcher, because indicators have to be justified also substantially and theoretically. However, when using statistical inference for making interpretations, the measurements from which these are drawn have to meet the required assumptions. In the case of measuring poverty and deprivation multidimensionally, the statistical assumption is that the indicators are quasi-parallel, i.e. that they are operationalisations of the same concept. To make sure that poverty and deprivation indicators meet this assumption, they have to pass the test of local independence.

Appendix A. Estimation of the LCM parameters

To obtaining maximum likelihood estimates for the LCM (in this case a LCM with three manifest A, B and C and one latent variable X) following equations A1-A3 have to satisfy equations A4-A7 (McCutcheon 1987, 21-27). Observed probabilities p_{ijk} are identified from latent ones π_{ijkt}^{ABCX} by using different notation. Conditional probabilities ($\pi_{it}^{\bar{A}X}, \pi_{jt}^{\bar{B}X}, \pi_{kt}^{\bar{C}X}$) indicate the probability that an observation in latent class *t* have also value, for example, *i* in variable A. Latent class probabilities π_t^X identify the number of latent classes and their relative sizes.

A1: $$\pi_{ijkt}^{ABCX} = \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_t^X$$

A2: $$\pi_{ijk} = \sum_t \pi_{ijkt}^{ABCX}$$

A3: $$\sum_i \pi_{it}^{\bar{A}X} = \sum_j \pi_{jt}^{\bar{B}X} = \sum_k \pi_{kt}^{\bar{C}X} = 1.00$$

A4: $$\pi_t^X = \sum_{ijk} p_{ijk} \pi_{ijkt}^{ABC\bar{X}}$$

A5: $$\pi_{it}^{\bar{A}X} = \frac{\sum_{jk} p_{ijk} \pi_{ijkt}^{ABC\bar{X}}}{\pi_t^X}$$

A6: $$\pi_{jt}^{\bar{B}X} = \frac{\sum_{ik} p_{ijk} \pi_{ijkt}^{ABC\bar{X}}}{\pi_t^X}$$

A7: $$\pi_{kt}^{\bar{C}X} = \frac{\sum_{ij} p_{ijk} \pi_{ijkt}^{ABC\bar{X}}}{\pi_t^X}$$

A widely used estimation method for obtaining maximum likelihood estimates to the LCM equations is the EM algorithm (Dempster et. al. 1977; Scott 1993). Its advance compared to other estimation methods is that it has two steps, where information gained in the first step is then used in the second step.

Because all latent class models here are unrestricted models i.e. all parameters are identified, the degree of freedom can be calculated using formula A8. The likelihood ratio chi-square for the goodness of fit test can be calculated using equation A9 and the test value follows X^2 distribution.

A8: $$df = (ijk - 1) - [(i + j + k - 2)t - 1]$$

A9: $$L^2 = \sum N\pi_{ijk} * Log(N\pi_{ijk} / F_{ijk})$$

Appendix B. Variables of the analysis

Name of the variable and explanation	*Coding*
A. Recipient of social assistance Household receive social assistance or housing allowance.	No => 0 Yes => 1
B. Problems in paying rent or mortgage Has the household been unable to pay scheduled rent for the accommodation / mortgage payments during the past 12 months?	No => 0 Yes => 1
C. Housing deprivation Accommodation have short of (1) space, (2) noise, (3) leaky roof, (4) damp walls, (5) rot in floors, (6) not enough light or (7) inadequate heating facilities. There is (8) pollution or (9) crime and vandalism in the area.	0 to 2 => 0 3 or more => 1
D. Relative deprivation Household does not have because it cannot afford: (1) car, (2) color TV, (3) video recorder, (4) micro wave, (5) dishwasher and (6) telephone.	0 to 1 => 0 2 or more => 1
E. Problems to make ends meet Is the household able to make ends meet? (1) With great difficulty, (2) with difficulty, (3) with some difficulty, (4) fairly easily, (5) easily and (6) very easily.	3,4,5 or 6 => 0 1 or 2 => 1
F. Relative income poverty Household's net income per consumption unit (OECD scale adjusted) less than 50% from the median of all households in the country.	No => 0 Yes => 1

REFERENCES

Alcock, Pete (1993): Understanding Poverty. MacMillan, London.

Atkinson, A. B. (1995): Incomes and the Welfare State. Essays on Britain and Europe. Cambridge University Press.

Atkinson, A. B. (1998). Poverty in Europe. Oxford: Blackwell Publishers.

Berghman, Jos (1995). 'Social exclusion in Europe: policy context and analytical framework in *Beyond the threshold: The measurement and analysis of social exclusion* (ed. Room, Graham), Bristol: The Policy Press.

Bollen, Kenneth A. (1989): Structural Equations with Latent Variables. John Wiley & Sons, Inc. New York.

Citro, Constance F. & Michael, Robert T. (1995): Measuring Poverty. A New Approach. National Academy Press, Washington D.C.

Dempster, A. P. & Laird N.M. & Rubin D. B. (1977): Maximin Likelihood from Incomplete Data via the EM Algorith. Journal of the Royal Statistical Society. Series B, Vol. 39:1, pp. 1-38.

Desai, Meghnad & Shah, Anup (1988): An Econometric Approach to the Measurement of Poverty. Oxford Economic Papers. Vol 40, pp. 505-522.

Eurostat (1999). ECHP Data Quality 108/99. Luxembourg: European Commission.

Gailly, Bernard & Hausman, Pierre (1980): Des Desavantages relatifs a une mesure objective de la pauvrete in Understanding Poverty (ed. Sarpellon, Giovanni).

Goodman, Leo A.(1978): Analyzing Qualitative/Categorical Data. Abt Books, Cambridge.

Creedy, John (1998): The Dynamics of Inequality and Poverty. Comparing Income Distributions. Edward Elgar, Cheltenham.

Heinen, Ton (1996): Latent Class and Discrete Latent Trait Models. Similarities and Differences. Saga, London.

Jenkins, Stephen P. (1991). 'Poverty Measurement and the Within-Household Distribution: Agenda for Action.' Journal of Social Policy, vol 20:4, pp. 457-483.

Kangas, O. and V. M. Ritakallio (1998). Different Methods - Different Results? Approaches to Multidimensional Poverty' in *Empirical Poverty Research in Comparative Perspective* (ed. Hans-Jurgen Andreβ), Aldershot: Ashgate.

Lazarsfeld, P.F. & Henry, N.W. (1968): Latent structure analysis. Houghton Mifflin, Boston.

McCutcheon, Allan L. (1987): Latent Class Analysis. Sage University Paper series on Quantitative Applications in the Social Sciences. Sage, Beverly Hills.

Moisio, Pasi (2001): The Nature of Social Exclusion – Spiral of Precariousness or Staristical Category? in Social Exclusion in European Welfare States (ed. Muffels, R. & Stakloglou, P.), Edward Elgar, Cheltenham. (Forthcoming)

Nolan, Brian & Whelan, Christopher T. (1996): Resources, Deprivation and Poverty. Clarendon Press, Oxford.

Ringen, Stein (1985): Toward a Third Stage in the Measurement of Poverty. Acta Sociologica 28:2, s. 99-11.

Ruggles, Patricia (1990): Drawing the line / Alternative poverty measures and their implications for public policy. Urban Institute Press, Washington.

Scott, Eliason R.(1993): Maximum Likelihood Estimation: Logic and Practice. Quantitative Applications in the Social Sciences N. 96. Sage.

Sen, Amartya (1979): Issues in the Measurement of Poverty. Scandinavian Journal of Economics 81, s. 285-307.

Townsend, Peter (1979): Poverty in United Kingdom. Penguin Books Ltd.. Middlesex.

Townsend, Peter (1987): Deprivation. Journal of Social Policy 16:2, s. 125-146.

Vermunt, Jeroen K. (1997): LEM: A general program for the analysis of categorical data. Tilburg University. The Netherlands.