Haute école de gestion de Genève
CRAG - Centre de Recherche Appliquée en Gestion
Cahier de recherche

A Cluster Analysis of Multidimensional Poverty in Switzerland

Giovanni FERRO LUZZI
Yves FLÜCKIGER
Sylvain WEBER

Cahier : N° HES-SO/HEG-GE/C--06/2/1--CH

2006

A Cluster Analysis of Multidimensional Poverty in Switzerland[*]

Giovanni FERRO LUZZI[§]
Yves FLÜCKIGER[§]
Sylvain WEBER[‡]

Cahier de recherche

juillet 2006

[*] This study has been realized using the data collected in the "Living in Switzerland Survey" (1999-2020) by the Swiss Household Panel (SHP), University of Neuchâtel. This project is financed by the Swiss National Science Foundation (Grants 5004-53205 / 5004-57894 / 5004-67304 / 10FI11-103293 / 10FI13-108500/1) (www.swisspanel.ch).

[§] University of Geneva, Switzerland.

[‡] Geneva School of Business Administration (HEG Genève), University of Applied Sciences of Western Switzerland (HES SO), e-mail: sylvain.weber@hesge.ch.

Résumé

La pauvreté est habituellement mesurée sur des bases financières uniquement. C'est cependant un état multidimensionnel, non seulement en rapport avec la situation financière, mais également avec la santé, la formation, l'environnement vital, l'état psychologique ainsi que la conjoncture sociale. Il faut par conséquent tenir compte de toutes ces composantes pour avoir une image complète de la pauvreté. Ceci est particulièrement vrai pour les pays industrialisés, où les personnes connaissant des difficultés financières peuvent bénéficier d'aides sociales telles qu'un revenu minimum ou des allocations de chômage et de logement. L'exclusion sociale et une mauvaise santé peuvent ainsi avoir un impact primordial dans le sentiment de pauvreté.

Dans cet article, nous proposons une méthodologie qui offre de nouvelles perspectives dans le contexte de la pauvreté multidimensionnelle. Dans une première étape, on effectue une analyse factorielle afin de construire des indicateurs de pauvreté basés sur de nombreuses dimensions potentielles et sans imposer de contrainte à priori. Les variables de base sont alors combinées dans quelques facteurs communs qui contiennent chacun une facette de la pauvreté multidimensionnelle. Une analyse typologique est ensuite réalisée sur la base des scores factoriels obtenus, dans le but de déterminer des sous-groupes de la population différemment affectés par les multiples dimensions de la pauvreté, ce qui permet d'identifier les pauvres. Une régression logistique est finalement estimée afin de rechercher les déterminants de la pauvreté.

Mots-clés

Pauvreté multidimensionnelle, Analyse factorielle, Analyse typologique, Suisse.

Summary

The measurement of poverty has often been criticized for relying solely on measures of financial deprivation. Poverty being a multidimensional state, related to health, schooling, living environment, psychological state as well as social tides, care should be taken to integrate these various components to have a proper picture of poverty. This is especially true for rich countries where poor financial conditions are often alleviated by social policies like minimum income, unemployment or housing benefits. Social exclusion and poor health can therefore dominate the poverty feeling.

We illustrate how some descriptive statistical tools can offer new insights in the context of multidimensional poverty. Factor analysis is used in a first step to construct poverty indicators based on many possible dimensions without posing too many a priori restrictions. The base variables are thus combined to produce common factors which convey some aspect of multidimensional poverty. By ascribing individual scores on each factor, we then use cluster analysis to determine population's subgroups that are unevenly affected by the various dimensions of poverty, what allows us to identify the poor. Finally, a logit regression is run to find the determinants of poverty.

Keywords

Multidimensional poverty, Factor analysis, Cluster analysis, Switzerland.

1. Introduction

In the literature, the basic notion that poverty should be measured on the basis of as large a number of components (attributes) as relevant and feasible has enjoyed increasing support. Since the seminal work of Townsend (1979) and others, it has been recognized that other aspects of life not necessarily related to income can impair human development, such as the access to public goods, health, or education. Many authors have come up with new approaches to provide poverty measures which account for its multidimensionality while maintaining desirable properties (Bourguignon and Chakravarty, 1999, 2003; Atkinson, 2003). One main conceptual issue is how to count multi-dimensional poverty. In other words, is multi-dimensional poverty the accumulation of deprivation in various components of what is considered "normal life" (the intersection approach) or should it be defined as the failure to access to one or more of the dimensions (the union approach)?

On the empirical side, a few studies have come out which aim at applying the idea of multidimensionality to the measurement of poverty. The UNDP human poverty index (HPI) is one such attempt which combines life expectancy, education, and health. This index, although widely used, has been criticized however for leaving out a monetary measure of poverty, while providing no clear basis for its weighting of the various sub-indices (see Bibi, 2002). The choice of what factors should be considered "poverty" or "deprivation" as well as the importance of each in the HPI has been said to be value laden.

The goal of this paper is threefold. One is to find a statistical tool that enables us to obtain a picture of poverty without too many a priori restrictions. Currently household panels contain variables that can be used to measure non-monetary poverty or deprivation. We shall describe how *factor analysis* can provide a meaningful description of poverty when many variables are possible candidates. This step leads to the construction of broader indicators of poverty which are common grounds to various subsets of variables, and which can be given a value for each individual.

The second goal of the paper is to identify the "poor" based on the newly constructed indicators of deprivations. To this end, we make use of another statistical tool, which is not very often favored by economists, namely *cluster analysis*. With this method, we aggregate individuals according to how similar they are with regard to their various scores of multiple deprivations. There, we attempt to see whether the "union" or "intersection" approach is more relevant to the data we use.

Finally, once we have identified the "multiply deprived", we use a complementary log-log model based on the panel dimension of the dataset to examine what affect the chances of being poor.

2. Pitfalls in the Measurement of Poverty

A large body of research exists on the measurement of monetary poverty. It is concerned with finding suitable indicators that best define poverty while satisfying every basic property they should. Recently, some authors have attempted to extend this literature to the multidimensional direction. It is not our purpose to discuss such literature here, but it is worth mentioning some difficulties inherent to the multidimensional approach.

2.1 Choosing the indicators of poverty or deprivation

It is not easy to determine what and how many indicators should be taken into account for measuring deprivation. There is an obvious trade-off between the possible redundancy caused by overlapping information and the risk of obviating some important variables. Further, once the idea that poverty is a multidimensional state has been accepted, one still has to define how far poverty should be measured in the various directions. Some authors may argue that it is essential to include indicators relating to standards of living or social relations, while other may have a more restrictive view on what needs should be included.

2.2 "Absence by choice" or deprivation

Another practical problem is related to the distinction between "absence by choice" and deprivation. Preferences can affect the consumption choice of a good, service or activity that may be judged as "necessary". Hence, an individual who does not have such goods or activities should be considered deprived only if she would consume them, could she afford them. This definition of deprivation can therefore only be used when the required information is collected in the household survey. One should however remain cautious in interpreting this kind of information, as the degree of value judgment can be very high and the subjective measure of own deprivation can be very different from one individual to another.

2.3 Aggregating and weighting the indicators

One further difficulty lies in the choice of aggregating or not the various indices of deprivation. Should they be averaged in a unique indicator of multidimensional poverty? This carries the clear advantage of summarizing the complexity of multiple dimensions in a simple way. On the other hand, such an aggregation causes a loss of information. Since a multidimensional phenomenon is studied, the search of a better description of such variety is an important goal.

One additional problem with the aggregation in a single index may arise when ascribing weights to the sub-indices of deprivation. Before aggregating indicators, it is indeed necessary to establish a weighting structure for each one given their different features. If each one is considered as a deprivation indicator with different importance, then the researcher must assign a different weight to each variable to reflect their differences. The first option is an equal weighting for each element. It is used in some papers as Townsend (1979). Alternatively, we can compute the weightings from the data. One possible strategy relies on a weighting structure based on frequencies, so that they are calculated as a function of the relative frequencies of the variables. For example, Halleröd (1994) and Deutsch and Silber (2005) give more importance to deprivation of goods considered as necessary by larger groups of the population. The importance of each indicator can be also computed by means of different multivariate statistical methods, as factorial analysis (Nolan and Whelan, 1996), principal components analysis (Ram, 1982; Maasoumi and Nickelsburg, 1988), or cluster analysis (Hirschberg et al., 1991).

2.4 Threshold definition

This step is related to the aim of any poverty or deprivation analysis: the identification of the poor population. The main problem that arises with any poverty analysis is the arbitrariness of the threshold choice. Moreover, defining a poverty line implies dividing the population in two distinct groups, “poor” and “non-poor”, which can be excessively restrictive in view of the multidimensional nature of poverty. Some authors (Cheli and Lemmi, 1995) have opted for an alternative methodology which relies on the concept of “fuzzy sets”. In this case, different degrees of deprivation are assumed instead of a dichotomy between poor and non-poor.

3. Methodology

The vast majority of empirical studies on poverty use some index of financial deprivation, be it the income or the consumption of a person or a household. There are obvious advantages to using a money metric to measure poverty, as it is quite easily interpretable, transparent and more or less comparable across countries.

We are here interested in a more descriptive approach of multidimensional poverty, which hopefully, can bring some insights into this topic and some answers to the pitfalls discussed in section 2 above. Typically, in the empirical literature, the various components of poverty are treated as separate dimensions. As stated in the introduction, our ambition is to see whether some criticisms made on a poverty indicator like the HPI can be met by letting somehow the data speak for themselves. The dimensions themselves will be selected on the basis of their relative importance in the data. The idea is very similar to that of Slottje (1991), who suggested that, when measuring the quality of life across countries, the indicators could be weighted by the variance of individual attributes. To this end, he uses the method of principal components analysis.

Factor analysis has been used before in the study of poverty. Nolan and Whelan (1996) use it to select the most appropriate indicators of deprivation. Halleröd (1995) does not exclude any indicator but varies the weights. Here, we follow Dekkers (2004) in that factor analysis is used on all variables pertaining to some kind of deprivation and let the data determine how many *latent factors* are to be used, as well as the weights imposed on them. The approach is similar in spirit to Dewilde (2004) too, who uses latent class analysis, a categorical variant of factor analysis. Finally, such an approach has also been used by Collicelli and Valerii (2000), who use principal component analysis on a set of developing countries.

The main idea to describe multidimensional poverty is based on the assumption that its various components translate into several variables, on which individuals accumulate deprivation. Each component therefore constitutes a given set of “capabilities”, be it financial conditions, housing environment, social interactions, health or any other state that may hinder human development. Financial deprivation may translate into failure to repay debts, sacrificing vacations or unhealthy food purchases. Housing deprivation would imply smaller rooms, absence of central heating, or noisy living environment. In other words, each measured variable x_j is due to some *unobserved* common factors F_k and an idiosyncratic effect s_j:

$$x_j = \sum_k a_{jk} f_k + s_j$$

or, in matrix notation: $$x = A \cdot f + s$$

Where the x vector includes all observed (standardized[1]) variables, A is the matrix of factor *loadings*, f is the vector of (latent) common factors, and s is the unique effects of the variables[2]. There are various methods to extract the factors, one simple way being through principal components[3], whereby the eigenvalues and eigenvectors of the correlation matrix on the observed variables are solved, with the added (scaling) constraint that the sum of the squared eigenvectors is equal to total variance.

One problem we must address is the fact that all of our deprivation indicators are ordinal with only a few scale steps or even dichotomous. In such instances, it is known that the Pearson's correlation matrix is biased and will unavoidably lead to biased estimates of the factor loadings if used as the basis for a factor analysis (Olsson, 1979). We will thus calculate the tetrachoric and polychoric correlation coefficients between our original indicators and use the resulting matrix as the starting point of our factor analysis. As Knol and Berger (1991) suggest, we will use unweighted least squares.

The first step of the factor analysis is to decide how many factors are relevant to the model. As we shall see in the empirical part, this choice is not always clear and brings straight away some difficulty.

The next problem encountered in the factor analysis is that the factor loadings matrix A defined above is not uniquely determined. To ensure a solution, one has to introduce constraints on the parameters in the original model. In general, one requires the first factor to have maximal contribution to the common variance of the observed variables, the second to have maximal contribution to this variance subject to being uncorrelated with the first, and so on. However, it is possible that a more interpretable solution can be achieved using a transformed model, obtained by a process known as factor rotation. Various methods for the rotation of factors are available and we will make use of an oblique one (promax with power 3), which allows the factors to be correlated, rather than independent. In our case, this is indeed what we want, as we expect the different dimensions of poverty to be linked: several deprivations are likely to occur simultaneously.

Once a representation of the data in this form is considered adequate, every individual can be ascribed a "score" on each derived factor. These scores inform us on how each individual performs on each dimension of poverty. As all variables have been normalized, they indicate this performance relative to the mean of the population which is zero. Moreover, as all variables have been designed such that a higher value corresponds to a worse situation (see Table A1), individuals with negative scores fare better than the average on these dimensions, while the opposite is true for positive scores. Obviously, with the four latent dimensions we found, some individuals may score negatively on some dimension and positively on some other.

We want to identify some groups in the population which are more or less homogenous when using these measures of multidimensional poverty. To this end, we rely on cluster analysis. The latter is a technique which allows the classification of similar objects into different groups, or more precisely, the partitioning of an original population into subsets (clusters), so that the data in each subset (ideally) share some common trait – often proximity according to some defined distance measure. The goal is thus to bring together individuals having relatively similar characteristics, while individuals belonging to different groups are as disparate as possible. With the agglomerative hierarchical clustering method we will use, the main steps of the groups' identification procedure are as follows. Let there be n individuals with m characteristics (in our case the various scores of poverty). At the beginning, every

1 Every variable that enters the factor analysis has to be standardized for two reasons. First, if the units of measurement are different across variables, they would not be comparable. Second, if there are large difference between the variances of the original variables, those whose variances are largest will tend to dominate the early factors.

2 The variables must be normalized as the procedure is sensitive to the units of measurement. Without this step, a variable with greater units could load higher on some factors, although the correlations are obviously the same.

3 See Everitt and Dunn (2001) for a detailed account of Factor Analysis.

individual is considered as a separate group. A similarity index, namely the Euclidean distance between the average scores of two clusters, is computed for all $n\cdot(n-1)/2$ potential pairs of individuals and the two closest are grouped. In the next step, the same procedure is applied to the (n-1) remaining clusters, which implies $(n-1)\cdot(n-2)/2$ distances. This process goes on until all observations belong to the same group, and hence creates a hierarchy of clusters.

The agglomerative hierarchical clustering method leaves open the choice of the final number of clusters. Many stopping rules can help this decision and we will make use of the best two among the thirty investigated by Milligan and Cooper (1985). If possible, the number of clusters will thus be chosen such that the information loss is limited (the number of clusters is set as the number where pseudo-t^2 is maximal plus one) while the difference between the clusters (the pseudo-F) is maximized.

Once the best groups based on the dimensions of poverty are established, we are able to identify one or several groups of poor. As stated, it is theoretically possible that many groups are formed where only some dimensions are relevant to poverty. In other words, poor can be found either in the "intersection" sense (poverty in all dimensions) or the "union" sense, if some deprivation is compensated by some non-deprivation on another dimension. Against this prospect, we should also keep in mind that, by construction from factor analysis, some dimensions are more relevant than others since they capture covariation in the deprivation variables in decreasing order.

As a result, the threshold of poverty itself will be defined by the observation of the various sub-groups of the original population. This flexibility in the definition and measurement of multidimensional poverty carries the advantage that no subjective choice needs to be made. On the other hand, such an approach clearly also has some drawbacks. The indicators of multidimensional poverty we use have little to no linkage to the axiomatic approach. Further, some subjectivity cannot be avoided with the statistical methods we use, as we will see. The pattern of deprivation and the relations among variables, especially in cluster analysis are not always clear cut, so that some choices must be made based on judgment, rather than on strictly statistical tools.

The final step of our analysis consists in finding the determinants of poverty. The group of poor individuals revealed by the cluster analysis is used and compared to the reference group. In order to find these determinants, we follow Dekkers (2004) by appealing to a simple panel logit model, where the dependent variable is the fact of being classified as poor or not (belonging to the group of poor or not). "Poverty" will be here defined by taking individuals who do not belong to the major cluster of "non-poor".

4. The Data

In this paper, we apply the technique presented in section 3 to Swiss data using the Swiss Household Panel (SHP). This panel dataset is very similar to the European Community Household Survey and consists of 5 waves, from 1999 to 2003. Questions are conceived to describe households as well as individuals and cover demographic, income, earnings, benefits, education, labor market status, description of housing and living conditions, possession of durables, mental and physical health, and so forth.

Table A1 in the Appendix is a list of the variables used for our factor analysis and Table 1 below provides basic statistics for these variables, all of which relating to some state of deprivation. As can be seen from Table A1, many of these variables describe situation of *financial deprivation.* However, the data set also includes interesting information on the state of health, the housing conditions, the environment, as well as variables pertaining to "social exclusion". In addition, we have also introduced "subjective" variables indicating the level of satisfaction with the financial situation or life in general, in the manner of Cheli and Lemmi (1995), Dekkers (2003, 2004) as well as Dewilde (2004). It may seem strange to include such variables as they do not reflect deprivation of some kind, but how this (specific or general) deprivation is felt by the individual. One may however argue that they may better reflect the everyday reality of poverty and deprivation.

In order to ease the interpretation of the coefficients obtained with the factor analysis and later with the cluster analysis, all variables have been constructed so that a greater value (or a value of one, if the variable is dichotomous) indicates a higher state of deprivation.

Table 1 gives the means, standard deviations, minimum and maximum value for each variable. In a headcount perspective, the mean deprivation levels vary quite substantially across binary ("have/have not") variables from a low 2% to almost 25%. If one looks at the variables close to what could be labeled "financial poverty", the range is somewhat narrower, extending from around 2% to 13%. Similar values are observed for the other waves of the panel.

Table 1: Descriptive statistics for the variables used in factor analysis, SHP 2001

Variable	Mean	Min	Max
Unpaid Bills	0.079	0	1
Cannot afford saving 100CHF	0.120	0	1
No private retirement scheme	0.099	0	1
Difficulty to make ends meet	2.582	0	10
Income below needs	0.099	0	1
Housing small	0.119	0	1
Bad heating	0.072	0	1
Cannot afford vacation	0.060	0	1
Cannot afford invite friends	0.030	0	1
Cannot afford restaurant	0.125	0	1
Cannot afford car	0.023	0	1
Cannot afford dishwasher	0.020	0	1
Cannot afford dentist	0.023	0	1
Cannot afford computer	0.026	0	1
*Financial satisfaction**	2.700	0	10
*HH financial satisfaction**	2.484	0	10
*Living standards satisfaction**	2.071	0	10
Noise in vicinity	0.209	0	1
Pollution in vicinity	0.145	0	1
Violence in vicinity	0.118	0	1
*Health status**	0.877	0	4
*Medication needed**	1.362	0	10
Handicap	0.186	0	1
Depression	1.867	0	10
*Life satisfaction**	1.892	0	10
*Optimism**	2.462	0	10
*Association membership**	0.247	0	1
*Cinema (frequency)**	2.947	0	4
*Sports (frequency)**	3.187	0	4
*Bar (frequency)**	1.748	0	4
*Theatre (frequency)**	3.047	0	4
*Contacts with friends (frequency)**	22.980	0	30
Observations		6'416	

Note: * The scale of these variables is inverted in comparison to what is usually done:
- For dummies: 0 = yes, 1 = no
- For frequency: 0 = every day to 4 = never (to 30 = never for *Contacts*)
- For satisfaction: 0 = completely satisfied to 10 = not at all satisfied

As explained in the text, the aim to define the variables in this manner is that a higher value in any variable will indicate a worse situation so that the interpretation of the following factor analysis will be eased.
A detailed description of the variables is available in Table A1.

Before analyzing the results of our multidimensional approach, we present evidence on the prevalence of poverty and its development over time in Switzerland based on a one-dimensional definition of poverty. Table A2 displays some basic results with respect to financial poverty using the most traditional indices such as the headcount ratio, the income poverty gap ratio or the FGT index (Foster, Greer and Thorbecke, 1984). These results have been computed for a poverty threshold fixed to 50% of the median equivalized household income. It is worth noting that the nominal relative poverty line has substantially increased in 2001 and 2002 whereas it slightly decreased in 2003.

Looking at Table A2, we observe that, in 1999, 7.8% of the entire population falls below the relative poverty line. This proportion stays more or less constant throughout the entire period analyzed despite the increase in the nominal threshold. A quite different picture arises when looking at the income gap ratio which shows that, in 1999, the poor were on average 22.1% away from the annual threshold value. In other words, in 1999 a transfer of 22.1% of 24'000 CHF was required to bring each poor to the limit of poverty. Four years later, the transfer needed to eradicate poverty has been reduced to 17.8% of the relative poverty line fixed at 25'500 CHF.

Looking at the different values taken by the FGT for higher values of α, reveals that the proportion of people living far away from the poverty line is very small. By increasing the poverty aversion from 2 to 5 reduces the FGT index from 0.626 to 0.089 in 1999. This conclusion is even more pronounced for 2003 which is characterized by an even more marked decrease of the FGT index.

These results confirm the conclusions obtained by a former study on poverty in Switzerland by Leu and Burri (1999), who furthermore show that poverty rates measured on a unique financial dimension are somewhat higher for women than for men, the difference however being virtually insignificant. They also highlight that poverty rates tend to be higher among single males and single parent households, among foreigners as well as in the French and Italian speaking regions. Finally, this study shows that poverty rates are very sensitive to the threshold definition.

5. Estimation of latent poverty factors and clusters of poor

Here we proceed with factor analysis in order to unravel what common factors best capture the covariance in all variables. As previously stated, all our variables are discrete. In such cases, tetrachoric and polychoric correlation coefficients must be used to estimate the factor loadings, as Pearson's correlation coefficients would lead to biased estimates. The polychoric correlation matrix itself is not presented as it is not our primary interest. Suffice is to say that, as expected, variables tend to be more strongly correlated when they belong to the same "dimension", although some exceptions can be observed. Only a few coefficients display a negative sign, but they are never significantly different from zero. This matrix of correlations is then used to extract the factors via unweighted least squares.

The next step involves choosing the appropriate number of latent factors. To this end, we rely on some standard visual and statistical tools, commonly used in factor analysis, although one should be aware that most of these rules are somehow *ad hoc* and cannot avoid value judgments. One method which has been put forth is to exclude factors with eigenvalues smaller than one, since the factors retained in this way account for more variance than the average for the variables. However this criterion is usually considered too lax, and should only be taken as an upper limit for the number of factors. Another method is to keep just enough factors so that the cumulated variance explained is no less than 70%. This criterion depends on the fit of the model. In our case, this would also imply a too large number of factors. The test we will finally use consists of an examination of the plot of the eigenvalues against the corresponding factor numbers, the so-called Scree diagram. The rate of decline tends to be fast for the first few factors but then levels off. The "elbow", or the point at which the curve bends, is considered to indicate the maximum number of factors to extract. One less factor than the number at the elbow might also be appropriate. Another way to use the Scree plot is to draw a straight line from the lowest eigenvalues, the threshold being where this line separates from the eigenvalues line.

The Figure A1 in the Appendix represents the Scree diagram for 2001. Only this year is provided, as the curves are very similar for the other ones[4]. In our case, the plot seems to indicate the presence of a general factor, as suggested by a large first eigenvalue (8.459) and a much smaller second one (2.523). But one might argue that a secondary elbow occurs at the fifth eigenvalue implying a four-factor solution. We thus decided that four factors were appropriate to describe the data, although the variance explained by the third and fourth factors is somewhat on the low side.

Next, we apply a rotation of factors to provide a more meaningful and easily interpretable solution loading matrix. As stated in the previous section, it makes sense to hypothesize that the common factors of deprivation (our four dimensions of poverty) are correlated. Indeed, one can assume that, say, "social exclusion" is positively correlated with "health" or "financial poverty". Therefore, we apply an *oblique* rotation that involves the introduction of correlations between factors. The resulting loadings for year 2001 are presented in Table 2. Once again, very similar results are found for the other years.

A glance at Table 2 shows some clearly distinctive patterns. Indeed, the first 13 variables load positively and quite high on the first factor. These variables all pertain to financial deprivation, or deprivations in basic goods and services that are due to the lack of financial resources, like sacrificing housing conditions, durable goods, and other activities usually taken for granted. It is worth noting that the three subjective indicators of satisfaction with the financial situation also have high loadings on this

[4] Table A3 in the Appendix contains the eigenvalues, as well as the associated proportion of variance explained by each latent factor for years 1999-2003.

Table 2: Rotated Factor Loadings (Oblique Rotation), 2001

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Unpaid Bills	0.673	0.072	-0.087	-0.165
Cannot afford saving 100CHF	0.816	-0.028	-0.064	0.087
No private retirement scheme	0.702	-0.070	0.058	0.023
Difficulty to make ends meet	0.788	0.062	-0.034	-0.112
Income below needs	0.518	0.001	0.041	0.151
Housing small	0.166	0.046	0.103	-0.153
Bad heating	0.269	0.147	0.039	-0.055
Cannot afford vacation	0.735	-0.019	-0.114	0.139
Cannot afford invite friends	0.726	-0.098	0.069	0.161
Cannot afford restaurant	0.717	-0.145	-0.006	0.143
Cannot afford car	0.583	0.016	0.140	0.090
Cannot afford dishwasher	0.556	-0.079	0.217	0.032
Cannot afford dentist	0.744	0.049	-0.036	-0.053
Cannot afford computer	0.715	0.030	0.043	0.112
Financial satisfaction	0.499	0.257	-0.005	-0.172
HH financial satisfaction	0.768	0.102	-0.022	-0.176
Living standards satisfaction	0.476	0.162	0.038	-0.154
Noise in vicinity	0.025	0.007	0.800	-0.049
Pollution in vicinity	0.006	-0.011	0.878	-0.008
Violence in vicinity	0.075	0.046	0.418	0.018
Health status	0.069	0.488	0.009	0.357
Medication needed	-0.032	0.400	0.010	0.508
Handicap	-0.007	0.384	0.004	0.347
Depression	0.050	0.624	0.006	0.027
Life satisfaction	0.246	0.542	0.009	-0.111
Optimism	-0.081	0.684	0.002	0.067
Association membership	0.181	0.050	-0.010	0.163
Cinema (frequency)	0.058	-0.043	-0.048	0.615
Sports (frequency)	-0.064	0.149	0.077	0.348
Bar (frequency)	0.142	-0.058	-0.015	0.330
Theatre (frequency)	0.252	-0.095	-0.039	0.235
Contacts with friends (frequency)	-0.095	0.053	-0.025	0.314

Inter-factor correlations

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1.000	0.323	0.226	0.032
Factor 2	0.323	1.000	0.198	–0.004
Factor 3	0.226	0.198	1.000	0.080
Factor 4	0.032	–0.004	0.080	1.000

Note: The *Promax* oblique rotation method has been used, with a power of 3.

factor. Hence factor 1 clearly reflects the dimension of "*Financial poverty*". Income tightness still reflects the main hidden factor of poverty in Switzerland.

The second factor is clearly related to physical (*Health, Medication* and *Handicap)* and mental (*Optimism*, *Depression* and *Life satisfaction*) health together. This latent dimension could be labeled "*Poor health*".

The next dimension which seems to have some importance could be named "*Bad neighborhood*". Only three variables (*Noise, Pollution* and *Violence*) loads pretty high on this factor.

Finally, the fourth factor has high loadings for variables that are mostly related to social life, like being member of an association, seeing friends or family, or simply going out. This latent factor is clearly associated to the dimension known as "*Social exclusion*". It is worth mentioning that the three variables relative to physical health also have pretty high loadings in this factor, indicating that they have a clear impact on social life.

Let us also mention that the third and fourth factors appear in reversed order for year 2003. This only means that for this particular year, the *"Social exclusion"* was more important than the *"Bad neighborhood"*. As shown in the columns "proportion" of Table A3, the shares of variance explained by these two factors are always close and we thus do not need to concern any longer about this inversion.

Dekkers (2004) only identifies three factors for Belgium, which are "Financial poverty", "Social exclusion" and "Poor *mental* health"[5]. Dewilde (2004) finds four underlying factors when analyzing British and Belgian households' panels. However, they are different from ours, as she lists "Housing", "Financial stress" and "Limited financial means" which jointly correspond roughly to our "*Financial poverty*" and "Housing environment", which is the same as our "*Bad neighborhood*". Nevertheless, it must be emphasized that all of these names are only subjective labels based on the examination of the loadings resulting from factor analysis and the rotation performed.

The second part of Table 2 gives the correlation coefficients among the four factors, as implied by the oblique rotation. It appears that factors 1, 2 and 3 are moderately and positively correlated, while factor 4 has no correlation with the other factors. It seems therefore that financial poverty, poor health and bad environment move together to some extent, whereas social exclusion is unrelated to the other dimensions. This last result is somewhat unexpected, as we had anticipated a strong positive correlation, at least with financial poverty, as found in Dekkers (2004).

There is however some inconsistency across the years for this factor. In 1999, "*Social exclusion*" is positively (but still moderately) correlated to all other dimensions except "*Bad neighborhood*", while it is slightly positively correlated only to "*Bad neighborhood*" in 2003. We therefore prefer to remain inconclusive as to the correlation of "*Social exclusion*" with the other dimensions of poverty.

We now turn to the results of the cluster analysis. As detailed in section 3, individuals are being grouped according to their relative distance (Euclidean distance), and the appropriate number of groups or "clusters" is determined by looking at various statistics. As is well-known, such an exercise is highly subject to value judgment. We considered two different statistics, namely: the pseudo-F developed by Calinski and Harabasz (1974), and the pseudo-t^2 which is a transformation of the Je(2)/Je(1) presented by Duda and Hart (1973). Large values of the pseudo-F index indicate distinct clustering and one must therefore maximize this statistic. The opposite is true for the pseudo-t^2, and one should choose the number of clusters so that this index is low and has much larger values next to it. It is advisable to look for a consensus among the two statistics, that is, local peak of the pseudo-F statistic combined with a small value of the pseudo-t^2 statistic and a larger value of the latter for the next cluster fusion.

Both of these statistics are displayed in Table 3, where the first 10 cluster groupings can be examined. Taking 1999 as an example, we see that the pseudo-F is maximized for 3 clusters, whereas the pseudo-t^2 is maximal for 8 groups, indicating the presence of 9 clusters. But notice that the pseudo-t^2 is also high for 2 groups, so that the solution of 3 clusters seems to be the best compromise. Applying the same reasoning to each year gives two clusters for 2000 and 2001, and four clusters for 2002 and 2003.

The dendrogram (or cluster tree) in Figure A2 of the Appendix presents graphical information concerning which observations are grouped together at various level of similarity. At the bottom of the dendrogram, each observation would be considered its own cluster. As one climbs up in the tree, observations are combined until all are grouped together, the height of the vertical lines indicating the similarity (or dissimilarity) of two groups. Creating 2 clusters tantamount to cutting the tree horizontally where it has only two branches. Since they are among the longest branches, it confirms that the two clusters we formed are actually very dissimilar.

Table 4 shows the average scores of the individuals pertaining to the various clusters found in each year. Typically, a first very large cluster contains most of the sample, and can undoubtedly be defined as the "non poor" cluster. The mean scores are found to be negative on all dimensions of poverty, indicating that most persons are not deprived in any direction. A smaller second cluster is then found to have positive mean scores on every dimension. The individuals belonging to this cluster can thus be called "multidimensional poor", since they suffer from multiple deprivations. For the years with more than two clusters, we see that the further groups can be considered as outliers, as very few individuals compose them. We finally obtain the following proportions of poor: 4.38% in 1999, 1.42% in 2000, 1.48% in 2001, 2.88% in 2002 and 2.98% in 2003.

[5] Dekkers (2004) only includes variables pertaining to mental health in his analysis, as he considers physical health to be a determinant of poverty, rather than a dimension of poverty itself. Although we admit that health could be taken as a determinant, we see no reason to separate mental and physical health.

Table 3: Statistics for determining the number of clusters

Number of clusters	1999 Pseudo-*F*	1999 Pseudo-t^2	2000 Pseudo-*F*	2000 Pseudo-t^2	2001 Pseudo-*F*	2001 Pseudo-t^2
1	-	27.61	-	574.06	-	506.23
2	27.61	1263.16	574.06	5.13	506.23	27.44
3	647.66	24.48	292.51	27.84	284.76	19.96
4	451.26	113.86	204.92	36.99	197.10	16.64
5	391.79	23.60	168.37	18.93	156.02	27.66
6	322.52	54.67	139.45	9.86	133.27	22.90
7	282.30	28.62	118.00	909.70	115.43	1997.17
8	247.32	1820.89	243.86	5.76	413.69	18.50
9	489.08	115.99	214.26	29.06	363.99	91.57
10	454.63	609.26	194.46	48.52	340.44	226.49

Number of clusters	2002 Pseudo-*F*	2002 Pseudo-t^2	2003 Pseudo-*F*	2003 Pseudo-t^2
1	-	68.67	-	56.16
2	68.67	90.49	56.16	57.21
3	80.11	604.58	56.99	575.27
4	260.95	143.4	233.96	47.60
5	235.36	68.17	197.54	78.41
6	211.59	10.92	176.02	13.31
7	179.48	6.75	149.81	2081.03
8	154.68	60.85	472.72	3.24
9	143.63	19.26	414.19	51.42
10	130.16	1653.05	377.2	40.39

Table 4: Mean Scores on the 4 Factors, by Cluster, 1999-2003

Year	Cluster	Factor 1	Factor 2	Factor 3	Factor 4	Observations	%
1999	1	-0.092	-0.074	-0.026	-0.040	7'397	95.58
	2	1.970	1.578	0.562	0.828	339	4.38
	3	3.374	2.777	1.557	3.429	2	0.03
2000	1	-0.050	-0.048	-0.015	-0.004	6'684	98.58
	2	2.829	2.182	0.506	0.013	96	1.42
2001	1	-0.049	-0.038	-0.012	-0.011	6'321	98.52
	2	2.863	1.885	0.564	0.217	95	1.48
2002	1	-0.076	-0.052	-0.031	-0.041	5'376	96.71
	2	2.338	1.147	0.682	1.008	160	2.88
	3	-0.291	2.634	2.195	1.217	17	0.31
	4	2.879	4.004	2.222	-0.148	6	0.11
2003	1	-0.089	-0.051	-0.016	-0.020	4'943	96.81
	2	2.575	1.213	0.387	0.574	152	2.98
	3	1.991	3.433	-1.712	0.148	8	0.16
	4	3.737	5.515	1.771	1.445	3	0.06

As one can see, the proportions we find are rather low. In fact, our measure is between 3 and 6% lower than the traditional headcount ratio, depending of the year. Different explanations could be offered: either the headcount ratio overestimates the number of poor people (or corollary our measure underestimates it), either the fact of taking into account several dimensions of poverty allows some sort of "compensation" between them. A financially poor individual according to the headcount ratio could effectively be classified in the non-poor by our method if he is doing as well as the average or so in the other dimensions of poverty ("*Poor health*", "*Bad neighborhood*" and "*Social exclusion*"). In some sense, this second reason is one of the goals of this method: we take into account more than just an income distribution, so that we measure something quite different and we do not obtain exactly similar results.

On the other hand, it is worth noting that the evolution of the various indices given in Table A2 is almost parallel with the development of our measure. Indeed, for almost every index, the value is high in 1999, decreases in 2000-2001, and increases in 2002-2003 but do not reach its level of 1999, which is exactly what happens with the proportions of poor we obtain.

6. The Determinants of Poverty

Our goal is here to assess the determinants of multidimensional poverty. The clusters we found enable us to build a dichotomous variable stating whether a person belongs to a group of poor or non-poor, for each year. One could in principle imagine that such a model be estimated as a multiple outcomes one, whereby individuals end up as “poor”, “partially poor” or “non–poor”, if one accept the “union” approach, in which being deprived along some but not all dimensions can also be considered as poverty. However, because the results of our cluster analysis did not provide such clusters in a consistent manner for each year and with sufficient observations in each group, it would not be very meaningful. One important step is the choice of the determinants of poverty. Clearly, all variables that are assumed to be potential causes of poverty should be included. We therefore selected variables that pertain to human capital as well as variables that may capture discrimination in the labor market, such as age, gender, and nationality (see Table A4). We also included variables that may be more causes of social exclusion such as household composition, marital status and the like. Finally, we introduce a set of time dummy variables to capture the effect of a given year on poverty. We did not include any measure of income as independent variable firstly because doing so would distort the coefficients estimated for the other covariates. Secondly, it would also bias the comparison between the model explaining the multidimensional poverty and the one explaining the financial poverty, since both would then not be based on exactly the same dependent variables. Practically, it is inconceivable to introduce the income as independent variable in the latter model because the dependent variable itself is constructed from the income, which would generate obvious problems of collinearity.

Our dependent variable being binary, we will use limited-dependent-variable models. Indeed, what we want to explain is the state of being poor, which can only be either true (1) or false (0). Several binary response models are available, such as probit, logit or complementary log-log. The latter is the most appropriate to analyze our data, since unlike the two others, it is asymmetric. This model is typically used when the positive outcome is rare, which is obviously our case with around 3% of poor individuals. Another desirable feature of the complementary log-log model is that it is the discrete-time equivalent of the Cox proportional hazard model which is widely used in duration analysis.

A glance at Table 5 reveals that most coefficients have the expected sign, though some variables appear to be insignificantly different from zero[6]. Gender, for instance, has no effect on poverty, per se, what corroborates the results of Dekkers (2003, 2004) for other European countries. On the other hand, the marital status dummy variables clearly indicate that divorced persons have a higher probability of falling into poverty with respect to both married and unmarried people. Also very much in line with expectations is the effect of education, which unambiguously lowers the chances of falling into poverty. Single parents also logically suffer more of multidimensional poverty.

Unemployment is also a strong predictor of poverty, as well as retirement. Being a foreigner raises the probability of belonging to the poor group, but differently with respect of the origin. Indeed, individuals coming from the “old” European Union at 15 (EU15) are less likely to be poor than other immigrants. Age does increase the chances of being poor in the first place and then decreases them (maximum at the age of 42-43). One could hypothesize that poverty increases with age, especially when households have children and therefore greater needs, and decreases progressively as children become less of a burden for their parents. This however does not square with the coefficients of the children variables. In this respect, children seem to play a mitigating effect on multidimensional poverty, at least when they are younger. It could be that the social exclusion factor is strongly reduced when households have children. Further, parents may prove choosier with respect to the quality of their environment, when they have children, which could further reduce the probability of belonging to the group of poor. Let us also mention that poverty is influenced by the linguistic regions of Switzerland. Multidimensional poverty is in fact more probable in the French-speaking part than anywhere else.

In order to have a broader picture of poverty, we ran the same model on a simple headcount indicator of financial poverty (with poverty line set at half the median income, see Table 6). It appears that the

[6] Because of non-linearity, the coefficient itself is only qualitatively related to the impact of the variable on the probability of falling into poverty. The latter is shown in the “marginal effect” column. For dummy variables, the effect is computed by taking the difference in probability, when the dummy is raised from 0 to 1, all other variables being set to their mean sample value.

Table 5: Complementary log-log model explaining multidimensional poverty

Variable	Coeff.	Std. Err.	Marginal Effect [a]	Std. Err.
Year 2000	-1.263***	0.131	-0.004960***	0.00048
Year 2001	-1.169***	0.131	-0.004605***	0.00047
Year 2002	-0.220**	0.107	-0.001081**	0.00050
Year 2003	-0.200*	0.110	-0.000989*	0.00051
Age/10	0.385*	0.214	0.002028*	0.00113
(Age/10)^2	-0.045**	0.022	-0.000239**	0.00012
Gender (female = 1)	-0.042	0.114	-0.000222	0.00060
Married	0.029	0.194	0.000152	0.00102
Divorced	0.739***	0.179	0.005245***	0.00170
Single parent family	0.776***	0.185	0.005874***	0.00195
Children < 18	-0.087	0.058	-0.000460	0.00031
Adu1t children 18-30	-0.165*	0.085	-0.000867*	0.00045
Adu1t children >30	0.039	0.246	0.000205	0.00130
Non-children adults	-0.450***	0.164	-0.002372***	0.00087
Education level 2	-0.337*	0.200	-0.001533*	0.00079
Education level 3	-0.869***	0.122	-0.004793***	0.00077
Education level 4	-1.581***	0.215	-0.005162***	0.00053
Education level 5	-1.752***	0.216	-0.005291***	0.00049
Part-time work	0.454***	0.148	0.002763***	0.00104
Student	-0.379	0.260	-0.001744*	0.00105
At home	0.625***	0.183	0.004264***	0.00159
Retired	0.961***	0.236	0.007504***	0.00265
Unemployed	2.371***	0.162	0.046927***	0.00807
Other occupation	0.517	0.367	0.003538	0.00319
EU15 national	0.747***	0.140	0.005476***	0.00140
Non-EU European	1.618***	0.272	0.020598***	0.00683
Other nationality	1.792***	0.330	0.025605***	0.00990
French Speaking Region	0.727***	0.222	0.004589***	0.00170
German Speaking Region	-0.273	0.220	-0.001512	0.00129
Intercept	-4.094***	0.599		
Ln σ^2_u	0.820	0.048		
σ_u	1.507	0.036		
Rho	0.580	0.012		
Log L	-2778.213			
Observations	27'487			
Groups (individuals)	8'575			

Notes: */**/*** Coefficient is significant at the 0.1/0.05/0.01 level. Reference category is "*Year 1999*", "*Male*", "*Single*", "*Education level 1*", "*Full-time work*", "*Swiss national*" and "*Italian speaking region*".
[a] Marginal effect evaluated at the mean of every variable and assuming that the random effect for that observation's panel is zero. For binary variables, variation of the probability of a "positive" outcome is calculated for a discrete change from 0 to 1.

estimation gives, broadly speaking, similar results. Variables like gender, single parent family, education, unemployment or retirement still have the same expected effect. It is nevertheless interesting to highlight the different effect of some variables, depending on which poverty is considered. For example, marital status seems to have no effect on financial poverty. Both coefficients on age become positive but not statistically different from zero, indicating that the effect of age upon financial poverty should be linearly positive. Age may thus have different effect, whether one looks at poverty in a strictly financial perspective, or more broadly defined on various indicators of deprivation. While the fact of being a student has no significant effect on multidimensional poverty, it clearly raises the chances of being financially poor. Having children (especially young ones) has a completely opposite effect on financial poverty. In a more traditional perspective, larger families have higher probabilities of belonging to the poor households' category. This effect could however be "exaggerated" because of the use of so called "expert" equivalence scales. In a recent paper, Falter (2005) has argued that such scales like the OECD scales are too steep i.e. such scales put too much emphasis on large families. Therefore, the impact of children on financial poverty is biased upward. He provides new estimates from income subjective data that show that equivalence scales are rather flat for families with three children or more. Thus, the supplementary income needed to support an

Table 6: Complementary log-log model explaining financial poverty (Equivalized Income less than the half of the median income)

Variable	Coeff.	Std. Err.	Marginal Effect [a]	Std. Err.
Year 2000	0.095	0.075	0.001365	0.00112
Year 2001	0.081	0.078	0.001166	0.00114
Year 2002	0.163 **	0.079	0.002406 *	0.00124
Year 2003	0.076	0.083	0.001092	0.00122
Age/10	0.164	0.165	0.002293	0.00231
(Age/10)^2	0.021	0.016	0.000290	0.00023
Gender (female = 1)	-0.003	0.092	-0.000046	0.00128
Married	0.012	0.170	0.000168	0.00237
Divorced	0.074	0.165	0.001060	0.00244
Single parent family	0.439 ***	0.159	0.007478 **	0.00328
Children < 18	0.615 ***	0.039	0.008609 ***	0.00064
Adu1t children 18-30	0.142 **	0.058	0.001993 **	0.00081
Adu1t children >30	0.359 **	0.177	0.005030 **	0.00248
Non-children adults	-0.784 ***	0.141	-0.010974 ***	0.00202
Education level 2	-0.116	0.160	-0.001543	0.00203
Education level 3	-0.715 ***	0.098	-0.010354 ***	0.00157
Education level 4	-1.492 ***	0.167	-0.013260 ***	0.00111
Education level 5	-1.878 ***	0.185	-0.014652 ***	0.00104
Part-time work	0.438 ***	0.129	0.007023 ***	0.00236
Student	1.051 ***	0.193	0.022603 ***	0.00604
At home	1.081 ***	0.141	0.024000 ***	0.00469
Retired	1.536 ***	0.180	0.041261 ***	0.00833
Unemployed	1.516 ***	0.178	0.046661 ***	0.01014
Other occupation	1.648 ***	0.228	0.055861 ***	0.01514
EU15 national	0.111	0.138	0.001628	0.00211
Non-EU European	1.186 ***	0.254	0.030820 ***	0.01099
Other nationality	1.797 ***	0.296	0.067042 ***	0.02280
French Speaking Region	-0.351 **	0.169	-0.004574 **	0.00207
German Speaking Region	-0.627 ***	0.162	-0.009932 ***	0.00295
Intercept	-4.009 ***	0.487		
Ln σ^2_u	1.097	0.040		
σ_u	1.730	0.034		
Rho	0.645	0.009		
Log L	-5478.180			
Observations		27'487		
Groups (individuals)		8'575		

Notes: See Table 5.

additional child is in fact less than that implied by the equivalized income we use in the present analysis.

Citizens from the EU15 are yet not statistically different from natives: only foreigners coming from outside the European Union at 15 have more chances of being poor than Swiss. Finally, the impact of living in the French or German part of Switzerland reduces the probability of being financially poor, maybe reflecting the fact that wages are higher in these regions than in the Italian part.

7. Concluding Comments

This paper has attempted to put forth some ideas to address well-known problems in the measurement of multidimensional poverty. The advantages of this approach can be summarized as follows. First, the number of dimensions as well as their relative importance is not determined *ex ante* but chosen on the basis of empirical regularities in the data. To this end, we have used factor analysis, although other statistical tools could be used alternatively. Their relevance is therefore directly dictated by their power in explaining the variance of various deprivation indicators, and we have found that such a method provides a parsimonious representation of multidimensional poverty.

Secondly, no poverty threshold needs to be set, since the population of multiply deprived persons is identified by looking at their similarities with respect to their scores on the various dimensions through cluster analysis. One concomitant advantage is that more than one group of poor can theoretically be identified, if clusters are found with different mean scores on the poverty dimensions. Based on the "union" approach of multidimensional poverty, some people could be identified as poor solely on some but not all dimensions. This evidently may call for different policy measures, depending on the degree of deprivation on each dimension. In our case, the clusters found showed only one relatively small group of poor, which would actually fit better with the "intersection" approach, since they were found to have positive mean scores on all dimensions.

Still, this approach also has some limits. No proper index of poverty aggregated over all dimensions may be computed, thus comparison is made difficult if one were to analyze different countries. Further, the statistical tools used (factor analysis and cluster analysis) may be subject to some arbitrariness, notably in the selection of the initial set of deprivation variables and in the choice of the number of groups retained.

Finally, our approach does not distinguish a possible sequence of multidimensional poverty. Do people fall into poverty sequentially in a similar fashion along the various dimensions, or do they become poor in no clearly distinguishing pattern? Such an issue should be addressed from an empirical point of view, as it may provide precious guidelines to policymakers concerned with poverty.

References

ATKINSON, A.B. (2003), "Multidimensional deprivation: contrasting social welfare and counting approaches", *Journal of Economic Inequality*, 1(1), pp. 51-65.

BIBI, S. (2002), "Mesures de la pauvreté dans une perspective multidimensionnelle: Une revue de la littérature", Mimeo.

BOURGUIGNON, F. and S.R. CHAKRAVARTY (1999), "A family of multidimensional poverty measures", in D.J. Slottje (ed.), *Advances in Econometrics, Income Distribution and Methodology of Science*, Essays in Honor of C. Dagum, Springer-Verlag, London.

BOURGUIGNON, F. and S.R. CHAKRAVARTY (2003), "The measurement of multidimensional poverty", *Journal of Economic Inequality*, 1(1), pp. 25-49.

CALINSKI, T. and J. HARABASZ (1974), "A dendrite method for cluster analysis", *Communications in statistics*, 3, pp. 1-27.

CHELI, B. and A. LEMMI (1995), "A totally fuzzy and relative approach to the multidimensional analysis of poverty", *Economic Notes*, 1, pp. 115-34.

COLLICELLI, C. and M. VALERII (2000), "A new methodology for comparative analysis of poverty in the Mediterranean: A model for differential analysis of poverty at a regional level", *Economic Research Forum*, Working Paper 2023.

DEKKERS, G. (2003), "Financial and multidimensional poverty in European countries: Can the former be used as a proxy of the latter?", IRISS working paper, No. 2003-13.

DEKKERS, G. (2004), La perception de la pauvreté face à la réalité. Mesure de la pauvreté multi-dimensionnelle d'après les données du PSBH. In : Doutrelepont, R, Mortelmand, D. Casman, M.-Th. (Eds) : *Onze Ans de Vie en Belgique. Analyses Socio-économiques à partir du Panel Démographie Familiale.* Gent : Academia Press. pp. 131-156.

DEUTSCH, J. and J. SILBER, (2005), "Measuring multidimensional poverty: An empirical comparison of various approaches", *Review of Income and Wealth*, 51(1), pp. 145-74.

DEWILDE, C. (2004), "The multidimensional measurement of poverty in Belgium and Britain: A categorical approach", *Social Indicator Research*, 68(3), pp. 331-69.

DRASGOW, F. (1969), "Polychoric and polyserial correlations", In Kotz, L. and N.L. Johnson (Eds.), *Encyclopaedia of Statistical Sciences*, Vol. 7, pp. 68-74, New York: Wiley, 1988.

DUDA, R.O. and P.E. HART (1973), "Pattern classification and scene analysis", New York: John Wiley & Sons.

EVERITT, B.S. and G. DUNN, (2001), "Applied multivariate data analysis", Edward Arnold, London.

FALTER, J.M. (2005), "Equivalence scales and subjective data in Switzerland", *Working Paper*, University of Geneva.

FOSTER, J., J. GREER and E. THORBECKE (1984), “A class of decomposable poverty measures”, *Econometrica*, 52(3), pp. 761-66.

HALLERÖD, B. (1995), “The truly poor: Direct and indirect consensual measurement of poverty in Sweden”, *European Journal of Social Policy*, 5(2), pp. 111-29.

HIRSCHBERG, J., E. MAASOUMI and D. SLOTTJE (1991), “Cluster analysis of measuring welfare and quality of life across countries”, *Journal of Econometrics*, 50(3), pp. 131-50.

KNOL, D. and M. BERGER (1991), “Empirical comparison between factors analysis and multidimensional item response models”, *Multivariate Behavioral Research*, 26(3), pp. 457-77.

KOLENIKOV, S. and G. ANGELES (2004), “The use of discrete data in PCA: Theory, simulations, and applications to socioeconomic indices”, *Working paper*, University of North Carolina.

LAYTE, R., B. MAÎTRE, B. NOLAN and C.T. WHELAN (2001), “Persistent and consistent poverty in the 1994 and 1995 waves of the European community household panel survey”, *Review of Income and Wealth*, 47(4), pp. 427-49.

LEU, R. and S. BURRI (1999), “Poverty in Switzerland”, *Swiss Journal of Economics and Statistics*, 135(3), pp. 303-28.

MAASOUMI, E., and G. NICKELSBURG (1988), “Multivariate measures of well-being and an analysis of inequality in the Michigan data”, *Journal of Business and Economic Statistics*, 6(3), pp. 326-34.

MILLIGAN, G.W. and M.C. COOPER (1985), “An examination of procedures for determining the number of clusters in a data set”, *Psychometrika*, 50(2), pp. 159-79.

NOLAN, B. and C. WHELAN (1996), “Measuring poverty using income and deprivation indicators: Alternative approaches”, *Journal of European Social Policy*, 6(3), pp. 225-40.

OLSSON, U. (1979), “Maximum likelihood estimation of the polychoric correlation coefficient”, *Psychometrika*, 44(4), pp. 443-60.

OLSSON, U., F. DRASGOW and N.J. DORANS (1982), “The polyserial correlation coefficient”, *Psychometrika*, 47(3), pp. 337-47.

RAM, R. (1982), “Composite indices of physical quality of life, basic needs fulfilment, and income. A ‘principal component’ representation”, *Journal of Development Economics*, 11(2), pp. 227-47.

SLOTTJE, D. (1991), “Measuring the quality of life across countries”, *Review of Economics and Statistics,* 73(4), pp. 648-754.

TOWNSEND, P. (1979), “Poverty in the United Kingdom”, Harmondsworth: Penguin.

TSUI, K. (2002), “Multidimensional poverty indices”, *Social Choice and Welfare*, 19(1), pp. 69-93.

Appendix

Table A1: Description of the variables used in the factor analysis, SHP

Variable	Label
Unpaid Bills	Bills unpaid in the last 12 months (1=yes, 0=no).
Cannot afford saving 100CHF	Cannot afford to save CHF 100 per month (1=cannot; 0=can).
No private retirement scheme	Cannot afford to apply for retirement saving scheme (1=cannot; 0=can).
Difficulty to make ends meet	Difficulties in making ends meet with current income (from 0="no difficulty" to 10="highest difficulty").
Income below needs	Household incomes are below necessary income (1=yes, 0=no).
Housing small	House or flat is too small (1=yes, 0=no).
Bad heating	Heating in house is bad (1=yes, 0=no).
Cannot afford vacation	Cannot afford one week of vacation (1=cannot; 0=can).
Cannot afford invite friends	Cannot afford to invite friends once a month (1=cannot, 0=can).
Cannot afford restaurant	Cannot afford restaurant once a month (1=cannot, 0=can).
Cannot afford car	Cannot afford a private car (1=cannot, 0=can).
Cannot afford dishwasher	Cannot afford a dishwasher (1=cannot, 0=can).
Cannot afford dentist	Cannot afford visit to the dentist if necessary (1=cannot, 0=can).
Cannot afford computer	Cannot afford a computer at home (1=cannot, 0=can).
Financial satisfaction	0=very satisfied with financial situation, 10= not at all.
HH financial satisfaction	0=very satisfied with household financial situation, 10=not at all.
Living standards satisfaction	Satisfaction with living standards (0=very satisfied, 10=not at all)
Noise in vicinity	Noisy environment (0=no, 1=yes)
Pollution in vicinity	Problems with polluted environment (0=no, 1=yes).
Violence in vicinity	Problems with delinquency or vandalism around the house (0=no, 1=yes)
Health status	State of health (0=very good, 4=very bad).
Medication needed	Needs of medication (0=no,10=very high).
Handicap	Frequency of negative feelings (0=never, 10=always)
Depression	Satisfaction with life in general (0=very satisfied, 10=not at all)
Life satisfaction	Optimism feeling frequency (0=always, 10=never)
Optimism	Long term health problem or disability of a psychological or physical nature (0=no, 1=yes)
Association membership	Passive or active member of whatever association (0=yes, 1=no)
Cinema (frequency)	Frequency of going to the cinema (0=every day, 1=at least once a week; 2=at least once a month, 3=less than once a month, 4=never)
Sports (frequency)	Frequency of going to sporting events (0=every day, 1=at least once a week; 2=at least once a month, 3=less than once a month, 4=never)
Bar (frequency)	Frequency of going to a bar, pub, restaurant (0=every day, 1=at least once a week; 2=at least once a month, 3=less than once a month, 4=never)
Theatre (frequency)	Frequency of going to the theatre(0=every day, 1=at least once a week; 2=at least once a month, 3=less than once a month, 4=never)
Contacts with friends (frequency)	Contacts with close friends per month (range is from 0=more than 30, to 30=no contact)

Figure A1: Scree Diagram for 2001 factor analysis

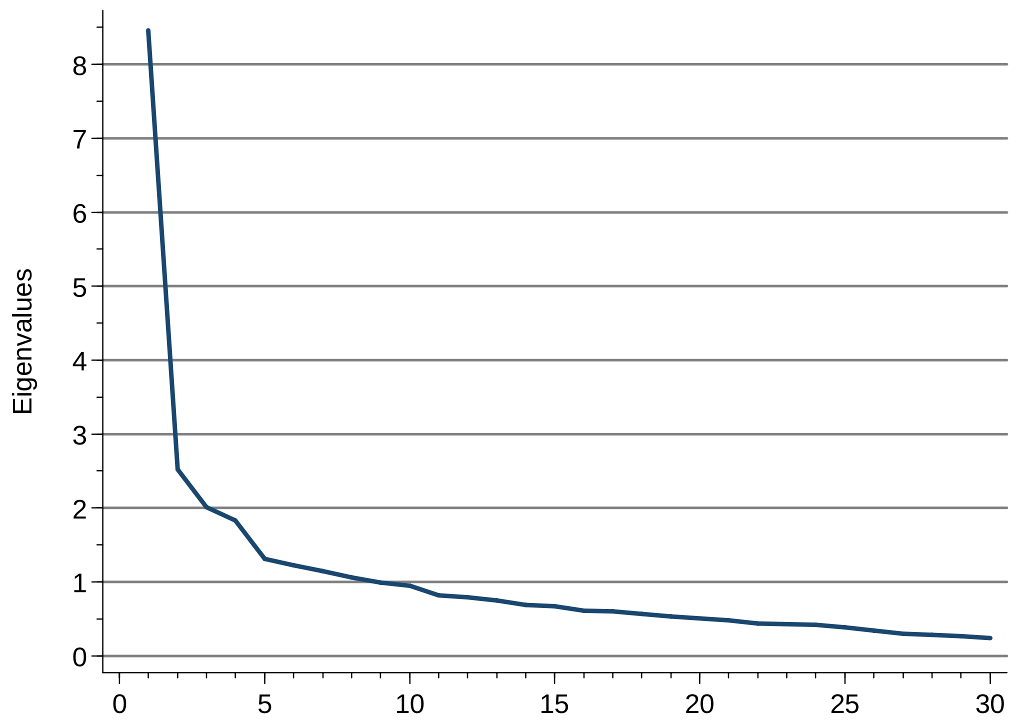

Table A2: Various Indices of Financial Poverty, 1999-2003

	1999	2000	2001	2002	2003
Poverty Line [a]	24'000	24'120	24'857	25'714	25'500
Headcount ratio %	7.787	7.533	7.329	7.895	7.523
Aggregate poverty gap [b]	412.960	295.830	279.980	360.720	340.760
Poverty gap ratio %	1.721	1.227	1.126	1.403	1.336
Income gap ratio %	22.096	16.281	15.369	17.769	17.763
Watts index	2.210	1.559	1.398	1.797	1.644
*Index FGT(0.5)*100*	3.352	2.542	2.495	3.007	2.820
*Index FGT(1.5)*100*	0.996	0.692	0.604	0.763	0.727
*Index FGT(2.0)*100*	0.626	0.428	0.360	0.467	0.430
*Index FGT(2.5)*100*	0.417	0.281	0.231	0.312	0.269
*Index FGT(3.0)*100*	0.289	0.194	0.156	0.223	0.175
*Index FGT(3.5)*100*	0.208	0.138	0.110	0.167	0.116
*Index FGT(4.0)*100*	0.153	0.102	0.080	0.131	0.079
*Index FGT(4.5)*100*	0.115	0.077	0.059	0.105	0.055
*Index FGT(5.0)*100*	0.089	0.059	0.045	0.087	0.039
*Clark et al. Index (0.10)*100*	2.149	1.517	1.365	1.745	1.607
*Clark et al. index (0.25)*100*	2.064	1.459	1.318	1.674	1.556
*Clark et al. index (0.50)*100*	1.935	1.372	1.247	1.570	1.476
*Clark et al. index (0.75)*100*	1.822	1.295	1.183	1.481	1.403
*Clark et al. index (0.90)*100*	1.760	1.253	1.148	1.433	1.362
*Sen index*100*	2.473	1.915	1.728	2.063	1.988
*Thon index*100*	3.366	2.412	2.214	2.747	2.621
*Takayama index*100*	1.674	1.201	1.100	1.363	1.302
Observations [c]	6’318	5’907	5’567	5’016	4’679

Notes:
[a] Poverty line is set at half the median equivalized income.
[b] Units of income per observation.
[c] The number of observations in not the same as in Table 1 because income was missing for some individuals.

Table A3: Eigenvalues and Proportion of Variance Explained

Factor	1999			2000			2001		
	Eigenvalue	Proportion	Cumulative	Eigenvalue	Proportion	Cumulative	Eigenvalue	Proportion	Cumulative
1	8.548	0.285	0.285	8.417	0.263	0.263	8.459	0.264	0.264
2	2.309	0.077	0.362	2.474	0.077	0.340	2.523	0.079	0.343
3	1.901	0.063	0.425	2.094	0.065	0.406	2.013	0.063	0.406
4	1.665	0.056	0.481	1.793	0.056	0.462	1.828	0.057	0.463
5	1.153	0.038	0.519	1.245	0.039	0.501	1.308	0.041	0.504
6	1.132	0.038	0.557	1.189	0.037	0.538	1.224	0.038	0.542
7	1.085	0.036	0.593	1.134	0.036	0.573	1.145	0.036	0.578
8	1.016	0.034	0.627	1.066	0.033	0.607	1.057	0.033	0.611
9	0.929	0.031	0.658	0.953	0.030	0.636	0.990	0.031	0.642
10	0.834	0.028	0.686	0.921	0.029	0.665	0.945	0.030	0.672
11	0.809	0.027	0.713	0.860	0.027	0.692	0.817	0.026	0.697
12	0.726	0.024	0.737	0.772	0.024	0.716	0.794	0.025	0.722
13	0.699	0.023	0.760	0.763	0.024	0.740	0.748	0.023	0.745
14	0.655	0.022	0.782	0.701	0.022	0.762	0.690	0.022	0.767
15	0.616	0.021	0.803	0.686	0.021	0.783	0.667	0.021	0.788
16	0.581	0.019	0.822	0.641	0.020	0.803	0.612	0.019	0.807

Factor	2002			2003		
	Eigenvalue	Proportion	Cumulative	Eigenvalue	Proportion	Cumulative
1	8.553	0.267	0.267	8.765	0.274	0.274
2	2.517	0.079	0.346	2.544	0.080	0.353
3	2.113	0.066	0.412	2.072	0.065	0.418
4	1.885	0.059	0.471	1.910	0.060	0.478
5	1.335	0.042	0.513	1.333	0.042	0.520
6	1.207	0.038	0.550	1.188	0.037	0.557
7	1.179	0.037	0.587	1.135	0.036	0.592
8	1.146	0.036	0.623	1.132	0.035	0.627
9	0.965	0.030	0.653	1.073	0.034	0.661
10	0.896	0.028	0.681	0.953	0.030	0.691
11	0.823	0.026	0.707	0.844	0.026	0.717
12	0.808	0.025	0.732	0.768	0.024	0.741
13	0.712	0.022	0.754	0.735	0.023	0.764
14	0.667	0.021	0.775	0.721	0.023	0.787
15	0.650	0.020	0.796	0.660	0.021	0.807
16	0.612	0.019	0.815	0.623	0.020	0.827

Figure A2: Dendrogram for 2001 cluster analysis

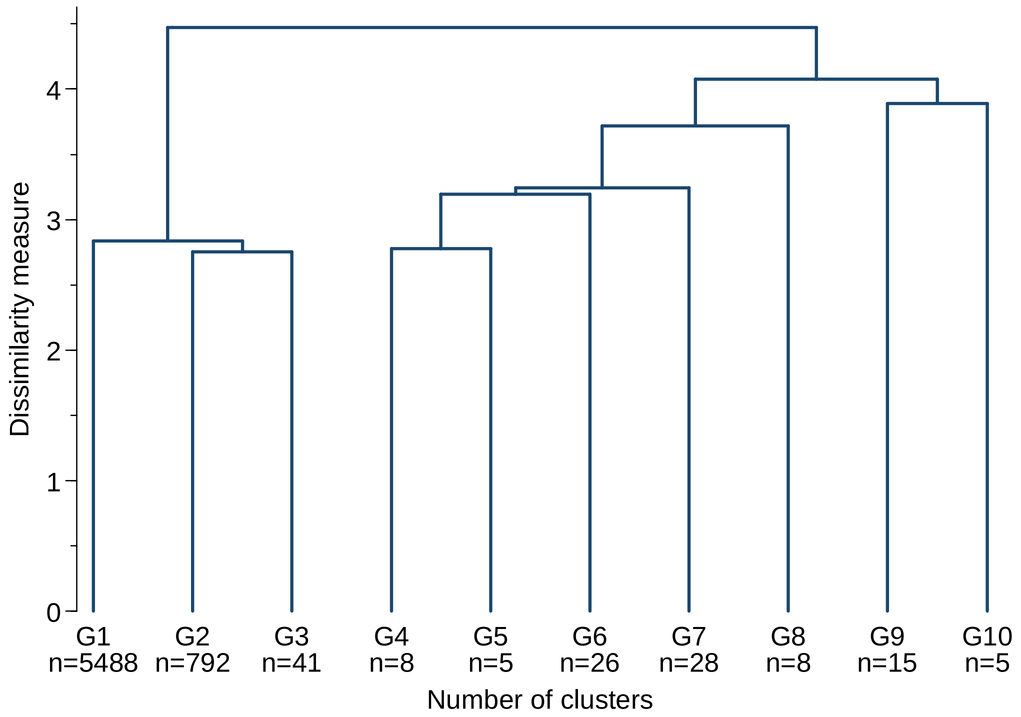

Table A4: Description of the variables used in the cloglog estimation, SHP

Variable	Label
Age/10	Age
(Age/10)^2	Age squared
Gender (female = 1)	Gender (0=male, 1=female)
Single	Civil status: single
Married	Civil status: married
Divorced	Civil status: divorced separated or widow
Single parent family	Parent living alone with one or more children
Children < 18	Number of children under 18 living in the household
Adu1t children 18-30	Number of adult children between 18 and 30 in the household
Adu1t children >30	Number of adult children over 30 in the household
Non-children adults	Number of non-children adults living in the household
Education level 1	Education: compulsory school or less
Education level 2	Education: domestic science/general training course
Education level 3	Education: maturity/apprenticeship
Education level 4	Education: technical/vocational school.
Education level 5	Education: university, higher specialized school
Full-time work	Occupation: fulltime job
Part-time work	Occupation: part-time job
Student	Occupation: student, apprentice
At home	Occupation: housekeeping
Retired	Occupation: retired
Unemployed	Occupation: unemployed or invalid insurance
Other occupation	Occupation: other
Swiss national	Nationality: Switzerland
EU15 national	Nationality: European Union 15
Non EU European	Nationality: Europe but outside EU15
Other nationality	Nationality: outside of Europe
French speaking region	French speaking region
Italian speaking region	Italian speaking region
German speaking region	German speaking region

Table A5: Descriptive statistics for the variables used in cloglog estimation, SHP 2001

Variable	Mean	Min	Max
Age/10	4.260	1.3	9.0
Gender (female = 1)	0.543	0	1
Single	0.295	0	1
Married	0.589	0	1
Divorced	0.116	0	1
Single parent family	0.060	0	1
Children < 18	0.864	0	6
Adu1t children 18-30	0.342	0	4
Adu1t children >30	0.019	0	2
Non-children adults	1.812	1	5
Education level 1	0.190	0	1
Education level 2	0.051	0	1
Education level 3	0.514	0	1
Education level 4	0.135	0	1
Education level 5	0.110	0	1
Full-time work	0.410	0	1
Part-time work	0.201	0	1
Student	0.131	0	1
At home	0.108	0	1
Retired	0.118	0	1
Unemployed	0.020	0	1
Other occupation	0.011	0	1
Swiss national	0.894	0	1
EU15 national	0.084	0	1
Non EU European	0.013	0	1
Other nationality	0.009	0	1
French speaking region	0.275	0	1
Italian speaking region	0.047	0	1
German speaking region	0.678	0	1
Observations		5'567	

Note: The number of observations is not the same as in Table 1 because income was missing for some individuals, who could therefore not be identified as financially poor or non poor. We did not take into account these individuals in order to estimate the two models (Tables 5 and 6) on exactly the same sample.

Cahiers de recherche du Centre de Recherche Appliquée en Gestion (CRAG) de la Haute Ecole de Gestion - Genève

CRAG – Centre de Recherche Appliquée en Gestion
Haute école de gestion - Genève
Campus de Battelle, Bâtiment F
7, route de Drize – 1227 Carouge – Suisse
✉ crag@hesge.ch
www.hesge.ch/heg/crag
✆ +41 22 388 18 18
🖷 +41 22 388 17 40

2006

- N° HES-SO/HEG-GE/C--06/1/1--CH
  Andrea BARANZINI
  Damien ROCHETTE
  “La demande de récréation pour un parc naturel
  Une application au Bois de Pfyn-Finges, Suisse”