

The Likelihood Ratio Test in Structural Equation Models: A Multi-Group Approach to Portuguese Material Deprivation

Paula C.R. Vicente

Assistant Professor, Department of Economics and Management, ULHT, Lisbon, Portugal; BRU-IUL, Lisbon, PORTUGAL

Corresponding Author: paula.vicente@ulusofona.pt

ABSTRACT

In this study a multi-group approach is used within the structural equations modelling framework, with the purpose of validating the temporal stability of the material deprivation measurement model, as well the impact of the existence of children in the household. For the purposes of model comparison, the use of the likelihood ratio test is discussed, when the normal distribution of the data cannot be assumed and the robust Satorra-Bentler scaled-corrected chi-square estimator is used. The statistical package LISREL 8.8 is used for estimating models.

Keywords— EU-SILC Data, Multi-Group Analysis, Satorra-Bentler Scaled Correct Chi-Square Estimator, Structural Equation Models.

I. INTRODUCTION

While the study of household poverty and quality of life is often based on disposable income, this measure alone may not be satisfactory. The level and quality of life of households can be measured through other indicators such as the ability to access a set of basic needs, the possession of durable goods, living conditions and even the environmental conditions of the place where they reside, that is, through a concept of material deprivation (Guio, 2005; Whelan & Maître, 2007).

This article is based on Portuguese cross-sectional data from European Statistical Database EU-SILC (European Statistics on Income and Living Conditions). EU-SILC is a four-year rotational panel, where each year 1/4 of the sample is recruited and 1/4 is dropped. Data from years 2007 and 2011 are used to analyze the material deprivation experienced by families, assuming this concept as measured in three dimensions: i) economic constraints;

(ii) possession of durable goods; and iii) housing conditions.

The aim of the study is to illustrate and discuss the use of the correction factor to the likelihood ratio test when the normality of the distribution of the data cannot be assumed. A multi-group structural equation model is used to validate the temporal stability of the material deprivation measure and the impact of the existence of children in the household on material deprivation. Statistical analysis was performed using the statistical package LISREL 8.8 (Jöreskog & Sörbom, 2004).

II. DATA

In this study the Portuguese EU-SILC cross-sectional data are used for the years 2007 and 2011, corresponding respectively to two independent samples with 4310 and 5740 aggregates. In fact, given the dynamics of panel rotation there is no overlap in respondents for the two years under analysis (see figure 1).

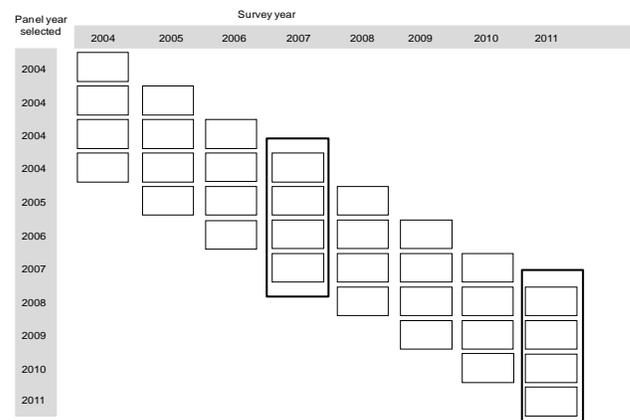


Figure 1: EU-SILC panel data design. In bold boxes are the data under analysis, 2007 and 2011

The variables under study were obtained as sums of items for each of the three dimensions of material deprivation considered: i) economic constraints; (ii) possession of durable goods; and iii) housing conditions.

The economic constraints dimension includes the lack of financial capacity to: i1) have a meal of meat or fish at least every two days, or the equivalent if vegetarian; i2) pay one week annual holidays away from home; i3) bear unexpected expenses, without resorting to credit; i4) keep the house adequately warm; and i5) meet the usual expenses and charges without difficulty. The possession dimension of durable goods considers the lack of economic availability to own: ii1) fixed or mobile telephone; ii2) color TV; ii3) washing machine; and ii4) car (passenger or mixed). The dimension housing conditions includes problems such as: iii1) leaking damp walls/floors/foundations, or rot in window frames or floor; iii2) insufficient light; iii3) absence of indoor toilet for sole use of the household; and iii4) absence of indoor bath or shower.

Thus, the economic constraints dimension takes values from 0 to 5 items of deprivation, the possession dimension of goods assumes values between 0 and 4 missing items, and the dimension housing conditions takes values between 0 and 4 problems. The distribution of deprivation observed in each of the three dimensions for the two years under analysis is shown in table 1.

Regarding the economic constraints, analysis of table 1 shows that 67.1% of the households do not have the financial capacity to meet 2 or more items in 2007, decreasing this percentage to 61% in 2011. Conversely, it should be noted that the high percentage of households that are not deprived of any item in possession of goods dimension, 82.7% and 85.8%, respectively, in 2007 and 2011. Also, the housing conditions dimension presents a high number of households with no problems, 67.2% and 73.5%, in 2007 and 2011, respectively.

When asked about the existence of children in the household, 36.83% of the families have children, in 2007, and that percentage is 31.8%, in 2011.

	Economic Constraints		Possession of Durable Goods		Housing Conditions	
	2007	2011	2007	2011	2007	2011
0 items	16.6%	20.0%	82.7%	85.8%	67.2%	73.5%
1 item	16.3%	19.0%	13.2%	11.7%	23.8%	21.2%
2 items	20.6%	22.4%	3.1%	1.7%	6.3%	3.8%
3 items	29.2%	24.4%	0.8%	0.6%	1.6%	1.1%
4 items	15.0%	12.1%	0.2%	0.2%	1.1%	0.4%
5 items	2.3%	2.1%				
	100%	100%	100%	100%	100%	100%

Table 1: Percentage of households with deprivation in each dimension (more items corresponds to more deprivation)

III. METHODOLOGICAL APPROACH

A structural equation model is used to explain the material deprivation, measured in the three dimensions mentioned above, as a function of the existence of children in the household, as represented in figure 2.

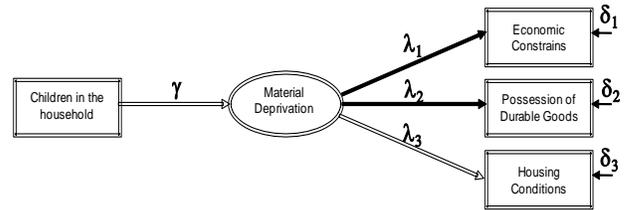


Figure 2: Structural equation model: having children is a possible determinant of material deprivation (measured by three indicators – the three material deprivation dimensions)

Structural equation modelling is one of the most important statistical techniques of multivariate analysis and can be seen as a combination of confirmatory factorial analysis (the measurement model) and linear regression (the structural model). The measurement model represents the relationship between the observed variables and the latent variables, while the structural model represents the dependencies among latent variables. It has the great advantage of allowing the simultaneous estimation of a set of dependency relationships between variables that may be latent, considering the component of the measurement error, Bollen (1989).

Regarding the model considered in this study and presented in figure 2, it should be noted that in terms of modeling, it is necessary to consider a latent variable underlying the observed explanatory variable, existence of children in the household, where γ the regression coefficient measures the relation between this latent variable and the latent material deprivation variable. The relationship between observed variables, economic constraints, possession of durable goods and housing conditions, with the latent variable material deprivation being measured by the factorial loadings represented by λ_1 , λ_2 and λ_3 , respectively. The measurement errors associated with the three observed variables are represented by δ_1 , δ_2 and δ_3 .

A longitudinal approach to the problem of material deprivation is proposed, using a multi-group structural equation model (Bollen, 1989; Schumacker & Lomax, 2010). In a multi-group analysis, it is possible to verify the existence of different behaviors by the different groups in relation to the concept under study, these groups being different temporal moments. For each of these groups a model must be defined. Then, equality constraints are imposed on parameters of the different groups.

Considered a baseline model (M_1), and a model in which restrictions are imposed to the parameters (M_0). M_0 can be considered a nested model of the baseline model, M_1 . From a statistical point of view, two nested models can be compared by calculating the difference between the chi-square values of the two models ($\Delta\chi^2$). Assuming data are normally distributed the difference is chi-square distributed. The corresponding degrees of freedom (Δdf) are given by the difference of degrees of freedom of the two nested models. The null hypothesis under test is the invariance of the parameters in both models, and the value of $\Delta\chi^2$ obtained must be compared with the critical value of a chi-square distribution with degrees of freedom given by Δdf . Rejecting the null hypothesis corresponds to concluding that the two models are not equivalent, that is, the baseline model (M_1) fits the data better than the restricted model (M_0). This chi-square difference test is also known as the likelihood ratio test.

IV. THE LIKELIHOOD RATIO TEST AND THE CORRECTION FACTOR

In structural equation models and for model estimation purposes, when the multivariate normality of the data distribution cannot be assumed, statistical packages such as LISREL and Mplus provide the use of a robust estimator or an ADF (Asymptotically Distribution Free) estimator: the Satorra-Bentler scaled-corrected chi-square. Such an estimator was proposed for data whose distribution deviates from the normal in terms of skewness and / or kurtosis (for metric variables or for ordinal variables for which a metric latent variable is assumed). When the normality of the data distribution cannot be assumed, and if the purpose is model comparison, the likelihood ratio test cannot be directly applied, since, due to the correction factor, the chi-square difference distribution is no longer a chi-square. In these circumstances, Bryant & Satorra (2012) recommend the application of a correction factor on the likelihood ratio test, for which an EXCEL macro for LISREL, EQS and Mplus users was made available in 2013, (Bryant & Satorra, 2013).

Using the adjustment measures provided by the statistical package, Min Fit Function Maximum Likelihood chi-square (T_1), Normal Theory Weighted Least Squares chi-square (T_2) and Satorra-Bentler chi-square (T_3), the procedure consists of two steps. First, a correction factor for the chi-square difference test (c_d) is calculated as follows

$$c_d = (df^{M_0}c_0 - df^{M_1}c_1)/(df^{M_0} - df^{M_1}), \quad (1)$$

in which, df^{M_0} and df^{M_1} are the degrees of freedom for the two models M_0 (model with imposed restrictions) and M_1 (baseline model), respectively. c_0 and c_1 are the correction factors for each of the models (M_0 and M_1) and are calculated as

$$c_0 = \frac{T_2^{M_0}}{T_3^{M_0}} \quad \text{and} \quad c_1 = \frac{T_2^{M_1}}{T_3^{M_1}}, \quad (2)$$

with $T_2^{M_0}$ and $T_3^{M_0}$ the obtained values of T_2 and T_3 , for M_0 model, $T_2^{M_1}$ and $T_3^{M_1}$ the obtained values of T_2 and T_3 for M_1 model.

Then Satorra-Bentler scaled difference chi-square test value (T_d) is computed as follows

$$T_d = (T_1^{M_0} - T_1^{M_1})/c_d, \quad (3)$$

with $T_1^{M_0}$ and $T_1^{M_1}$ the obtained Min Fit Function Maximum Likelihood chi-square values for the M_0 and M_1 models, respectively. The degrees of freedom for this corrected chi-square difference test are ($df^{M_0} - df^{M_1}$).

V. RESULTS

This study aims to illustrate the use of the macro built by Bryant and Satorra to estimate the proposed model to explain material deprivation in the years 2007 and 2011, using a multi-group approach. Thus, for each of the years 2007 and 2011, the model represented in figure 2 is specified, and a multi-group analysis was carried out to assess the temporal stability of the proposed factor structure, as well as the stability of the impact of the existence of children in the household at each of the two time points.

The models considered for comparison in multi-group analysis are

M_1 : Baseline model, with no restrictions of invariance between the two groups (subpopulations in 2007 and 2011);

$M_0^{(1)}$: Model with factor loadings invariance restriction in the two groups, that is, $\lambda_j^{2007} = \lambda_j^{2011}$, $j = 1,2,3$;

$M_0^{(2)}$: Model with regression coefficient invariance restriction, that is, $\gamma^{2007} = \gamma^{2011}$;

$M_0^{(3)}$: Model with factor loadings and regression coefficient invariance restrictions, that is, $\lambda_j^{2007} = \lambda_j^{2011}$, $j = 1,2,3$ and $\gamma^{2007} = \gamma^{2011}$.

Adjustment measures calculated by LISREL 8.8 statistical package for the four models under analysis are shown in Table 2. These results allow to calculate the chi-square test difference values for each of the three hypotheses, $H_0^{(1)}$, $H_0^{(2)}$ and $H_0^{(3)}$, in order to measure the invariance of the weights of different dimensions of material deprivation at the two time points, as well as the temporal invariance of the effect of the existence of children in the household on the material deprivation level.

Models	degrees of freedom	T_1	T_2	T_3
M_1	4	256.98	246.57	72.16
$M_0^{(1)}$	6	273.64	262.81	84.00
$M_0^{(2)}$	5	261.35	251.25	78.53
$M_0^{(3)}$	7	278.21	268.74	88.75

Table 2: Adjustment measures given by LISREL 8.8, for the models under analysis

$H_0^{(1)}$: $M_0^{(1)}$ is valid, that is, exists factor loadings invariance at the two time points, 2007 and 2011;

$H_0^{(2)}$: $M_0^{(2)}$ is valid, that is, exists coefficient regression invariance at the two time points, 2007 and 2011;

$H_0^{(3)}$: $M_0^{(3)}$ is valid, that is, exists factor loadings and coefficient regression invariance at the two time points, 2007 and 2011.

Table 3 presents the test values for the usual chi-square difference (assuming normal distribution of data) and for this difference with correction factor. The degrees of freedom of each one of the tests carried out are presented in the first column. The second and third columns correspond to the results of tests with no correction factor. The fourth column corresponds to chi-square tests with correction factor. The analysis of this table allows identical conclusions when comparing the models $M_0^{(1)}$ vs M_1 , and $M_0^{(3)}$ vs M_1 , using the chi-square difference test without the correction factor and with the correction factor. That is, the hypotheses of the invariance of the factorial loadings in the first case and, in the second case, of invariance of the factorial loadings and the regression coefficient in the years 2007 and 2011 are rejected. Consequently, M_1 must be chosen instead of $M_0^{(1)}$ and of $M_0^{(3)}$.

When comparing $M_0^{(2)}$ versus M_1 (that is, testing the hypothesis, $H_0^{(2)}$, that $M_0^{(2)}$ is valid), using the chi-square difference test without the correction factor (procedure considered incorrect), it is possible to conclude about the rejection of the hypothesis of regression coefficient invariance at the two different time points (2007 and 2011) and the consequent choice of M_1 . Using the chi-square difference test with the correction factor (procedure considered appropriate) the opposite conclusion is obtained. Thus, it is possible to conclude that there is no significant difference in the value of γ at the two different moments. In fact, the null hypothesis that $M_0^{(2)}$ is valid, and must be the chosen model, is not rejected.

	Δ degrees of freedom	Δ Min Fit ML χ^2	Δ Satorra- Bentler χ^2	Δ Satorra- Bentler scaled corrected χ^2
$M_0^{(1)}$ vs M_1	2	16.66 (0.0001)	11.84 (0.0013)	6.528 (0.0382)
$M_0^{(2)}$ vs M_1	1	4.37 (0.0215)	6.37 (0.0065)	1.8763 (0.1708)
$M_0^{(3)}$ vs M_1	3	21.23 (0.0000)	16.59 (0.0004)	9.4596 (0.0374)

Table 3: Obtained values of the chi-square difference (p values) and degrees of freedom

The factorial loadings of the indicators obtained in the model $M_0^{(2)}$ adjustment are presented in table 4, and it is possible to conclude that the economic constraints and possession of goods are the indicators that present greater weight in material deprivation in the years 2007 and 2011. The possession of goods dimension had an increasing relative weight from 2007 to 2011 in terms of material deprivation (from 0.78 to 0.83), while the relative weight of the housing condition dimension decreased in the same period (from 0.53 to 0.45).

	2007	2011
λ_1	0.71	0.71
λ_2	0.78	0.83
λ_3	0.53	0.45
γ	-0.14	

Table 4: Obtained values for structural equation model parameters, in 2007 and 2011

The estimated value for the impact of children in the household on material deprivation is negative, -0.14, which means that households with children have lower levels of material deprivation.

VI. CONCLUSION

The application of the multi-group analysis within the structural equation modelling framework allows testing the stability of the factorial structure of the material deprivation concept at the two time periods, 2007 and 2011, and / or the stability of the impact of the existence of children in the material deprivation level experienced by the families.

Thus, it is possible to conclude that there is no stability of the material deprivation measurement model in the two time periods, 2007 and 2011, increasing the weight of the possession of goods and decreasing the problems in housing conditions. On the contrary, there is evidence that the impact of the existence of children in the household on material deprivation levels remains constant from 2007 to 2011.

On the other hand, it is possible to conclude, that performing a multi-group analysis ignoring the correction factor on the statistic of the likelihood ratio test (and not using the Satorra-Bentler scaled chi-square difference test) may lead to statistically incorrect results. In some cases, these results can be contradictory.

ACKNOWLEDGMENT

The author acknowledges the support of Maria de Fátima Salgueiro, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal.

REFERENCES

- [1] Bollen, K. (1989). *Structural equations with latent variables*. New York: Wiley.
- [2] Bryant, F. B. & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. *Structural Equation Modelling*, 19(3), 372-398.
- [3] Bryant, F. B. & Satorra, A. (2013). *EXCEL macro file for conducting scaled difference chi-square tests via LISREL 8, LISREL 9, EQS and Mplus*. Retrieved from: www.eco.upf.edu/~satorra.
- [4] Guio, A. (2005) Material deprivation in the EU. *Statistics in Focus, Population and Social Conditions and Welfare*, 21.
- [5] Jöreskog, K.G. & Sörbom, D. (2004). *LISREL8*, Chicago IL: Scientific Software International, Inc.
- [6] Schumacker, R.E. & Lomax, R.G. (2010). *A beginner's guide to structural equation modelling*. (3rd ed.). Routledge, New York: Taylor & Francis Group.
- [7] Whelan, C.T. & Maître, B. (2007). Measuring material deprivation with EU-SILC: Lessons from the Irish survey. *European Societies*, 9, 147-173.