MEASUREMENT INVARIANCE: A METHOD BASED ON LATENT MARKOV MODELS

Francesco Dotto[1], Roberto Di Mari[2], Alessio Farcomeni[3] and Antonio Punzo[2]

[1] Department of Economics, University of Roma Tre, (e-mail: francesco.dotto@uniroma3.it

[2] Department of Economics and Business, University of Catania, (e-mail: roberto.dimari@unict.it), (e-mail: antonio.punzo@unict.it)

[3] Department of Economics and Finance, University of Tor Vergata, (e-mail: alessio.farcomeni@uniroma2.it)

ABSTRACT: We define differential item functioning in the context of panel data. We then present a general approach to detect measurement non-invariance cases in this context. We use a model selection procedure based on the Bayesian information criterion (BIC). A real data application and a simulation study are presented to illustrate and motivate the methods.

KEYWORDS: latent Markov model, measurement invariance, panel data

1 Introduction

Much empirical work in general social sciences leverages on questionnaire data analysis to measure possibly unobserved (latent) traits. In these contexts crucial working assumptions are that items have the same discriminatory power, unidimensionality of the latent trait, and measurement equivalence (invariance) in the scale.We focus on the assessment of potential violations of measurement equivalence in longitudinal studies. Namely, when respondents are repeatedly measured over time, and the model for the latent trait is not strong enough to describe the dependence structure among the items and external variables. This phenomenon is also known as differential item functioning (DIF), and we study it in connection with latent Markov models (see, e.g., Bartolucci *et al.*, 2013). We specify distinct notions of DIF, combining and extending ideas from Kankaraš *et al.*, 2018 and Masyn, 2017. Effectively, we develop a toolkit based on a classical model selection tool - i.e., the Bayesian Information Criterion - to select the most appropriate DIF configuration. An extended presentation of the technical framework, and of both numerical and real-data results is available in Di Mari *et al.*, 2022.

2 Mathematical Formulation

Let Y_{ith}, $h = 1,\dots,H$, be the h-th dichotomous indicator, measured for the i-th subject, $i = 1,\dots,n$, at time t, $t = 1,\dots,T$; observed alongside a vector of time-specific covariates $\mathbf{X}_{it}$. In addition, let U_{it} denote a discrete latent variable with support $\{1,\dots,K\}$, which follows a possibly inhomogeneous first-order Markov chain. In case of measurement invariance (no DIF), we assume the data arise from the following model

$$\begin{cases} P(Y_{it1} = y_1,\dots,Y_{itH} = y_H \mid U_{it} = k) = \prod_{h=1}^{H} \phi_{h|k}^{y_h}(1-\phi_{h|k})^{1-y_h}, \\ \log\left[\dfrac{P(U_{i1} = k \mid X_{i1})}{P(U_{i1} = 1 \mid \mathbf{X}_{i1})}\right] = \alpha_{1k} + \beta_{1k}\mathbf{X}_{i1}, \\ \log\left[\dfrac{P(U_{it} = k \mid U_{i,t-1} = j, \mathbf{X}_{it})}{P(U_{it} = j \mid U_{i,t-1} = j, \mathbf{X}_{it})}\right] = \alpha_{kj} + \beta_{kj}\mathbf{X}_{it}, \end{cases} \tag{1}$$

where the first equation denotes the measurement model which involves the item specific probabilities $\phi_{h|k}^{y_h}$. The two remaining equations define structural models for the initial and transition probabilities. The parameters α and β model the effect of the covariates on both the initial and transition probabilities. For the sake of simplicity we assume, as commonly done within this context, that such regression coefficients are time constant letting the covariate values be the driver of time heterogeneity. If DIF is allowed, the measurement model depends on $\mathbf{X}_{it}$ as

$$\text{logit}(\phi_{h|k}) = \gamma_{hk} + \eta_{htk}\mathbf{X}_{it}. \tag{2}$$

where γ is the intercept term and η_{htk} represents the direct effect of the covariate on the item specific probabilities. From equation (2) other DIF scenarios can be derived:

1. No DIF: The covariates only affect transition probabilities but they do not affect item specific probabilities.
2. Full DIF: The covariates affect both the transition probabilities and the item specific probabilities. The η_{htk} vector varies across items, time, and class.
3. Time-Constant DIF: The η_{htk} vector varies across items and class, but remains fixed across time.
4. State-Constant DIF: The η_{htk} vector varies across time and item, but remains fixed across latent states.
5. State- and Time- constant DIF: The η_{htk} vector is homogeneous across time and latent states.

3 Results

We analyse show syntetic simulation results and a real data analysis. These somewhat summarize the results reported in Di Mari *et al.*, 2022.

3.1 A simulation study

Table 1 reports the performance of the methodology in terms of rate of correct classification over 500 replicates for each setting. The fabricated data sets are based on $n = 500$, $T = 4$, $K = 3$, $H = 10$, with a single standard Gaussian covariate. It can be seen that the proper model is always selected with high probability.

3.2 General social survey: Measuring tolerance toward non-conformity

Data are taken from the American General Social Survey (GSS), a survey of the English-speaking, non-institutionalized adult population of the United States. The $H = 5$ binary items are formulated as follows: "Suppose ... wanted to make a speech in your community. Should he be allowed to speak?" and are referred to communists, atheists, militarists, homosexuals, and racists.

We include the covariate "Education", which we re-code into three categories. The best fit reveals a direct effect of Education on items, pointing out that to a higher education corresponds, on average, a higher probability to allow "Atheists" "Communists" "Homosexuals" "Militarists" and "Muslims" to speak in public.

The lowest BIC is attained at time- and state-constant DIF (DIF 4), i.e., for differing levels of education, individuals have varying probabilities of scoring "Yes" to the items, regardless of the underlying tolerance (latent) type, and record time (see Figure 1).

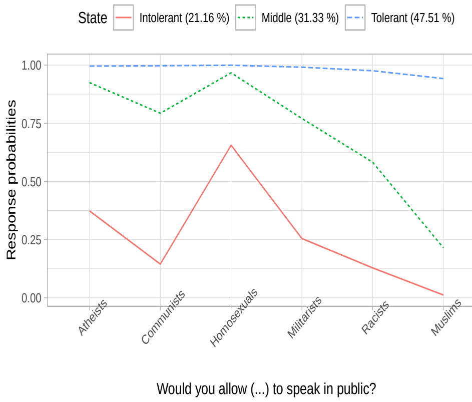
References

BARTOLUCCI, F., FARCOMENI, A., & PENNONI, F. 2013. *Latent Markov models for longitudinal data*. Chapman and Hall / CRC Press, New York.

DI MARI, R., DOTTO, F., FARCOMENI, A., & PUNZO, A. 2022. Assessing measurement invariance for longitudinal data through latent Markov models. *Structural Equation Modeling: A Multidisciplinary Journal*, **29**, 381–393.

BIC	True Model No DIF	 Full DIF	 Time Constant DIF	 State Constant DIF	 State Time Constant DIF
No DIF	1.00	0.00	0.01	0.00	0.00
Full DIF	0.00	1.00	0.00	0.00	0.00
Time constant DIF	0.00	0.00	0.99	0.00	0.00
State constant DIF	0.00	0.00	0.00	1.00	0.00
State Time constant DIF	0.00	0.00	0.00	0.00	1.00

Table 1. *Confusion matrix normalized by column to evaluate the BIC performance.*



	DIF 4
$\eta_{1\cdot\cdot}$	0.76***
("Atheists")	(0.17)
$\eta_{2\cdot\cdot}$	1.12***
("Communists")	(0.20)
$\eta_{3\cdot\cdot}$	0.85***
("Homosexuals")	(0.15)
$\eta_{4\cdot\cdot}$	0.39***
("Militarists")	(0.14)
$\eta_{5\cdot\cdot}$	0.09
("Racists")	(0.14)
$\eta_{6\cdot\cdot}$	1.16***
("Muslims")	(0.23)

Figure 1. *GSS data: Estimated response probabilities to answer "Yes" given state membership (on the left) and estimates of the direct effect η_h of the covariate "Education" on the six items available according to the time- state-constant DIF (on the right). Standard errors in parentheses are based on the observed Information matrix.*

KANKARAŠ, MILOŠ, MOORS, GUY, & VERMUNT, JEROEN K. 2018. Testing for measurement invariance with latent class analysis. *Pages 393–419 of: Cross-Cultural Analysis*. Routledge.

MASYN, K. E. 2017. Measurement invariance and differential item functioning in latent class analysis with stepwise multiple indicator multiple cause modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, **24**, 180–197.