# EPC-interest for categorical data with an application to the Inglehart values ranking

DL Oberski JK Vermunt G Moors

Department of Methodology and Statistics

Tilburg University, The Netherlands

ABSTRACT

Social science is often concerned with comparing groups such as countries, regions, time periods, the genders, or ethnicities. Because such substantive comparisons may be threatened when measurement parameters differ across the groups, it is common practice to test the hypothesis of exact equality of measurement parameters across groups: "measurement invariance testing". However, not all measurement differences are substantively relevant. At the same time, it was recently shown that not all substantively relevant differences are necessarily detected by the current best practices of invariance testing. Therefore, it was recently suggested to detect the relevant set of violations of measurement invariance using sensitivity analysis. This approach, recently introduced in this journal, uses the "EPC-interest", a measure of the change in the parameters of substantive interest due to violations of invariance. So far, however, EPC-interest was limited to continuous linear structural equation models and did not consider the impact of freeing several restrictions at the same time.

Since categorical data are common in the social sciences and measurement invariance restrictions may be highly correlated, this paper extends the "EPC-interest" to models with categorical observed and latent variables. Moreover, we explicitly consider freeing (potentially large) sets of equality restrictions simultaneously. The advantages of the extended EPC-interest are demonstrated using a 48-country analysis of "materialism / postmaterialism" values measured by three partial ranking tasks using multilevel latent class regression. Some corroboration for hypotheses on postmaterialism from the literature are found when employing the EPC-interest, while without its application some of the findings contradict these hypotheses.

The newly developed methods discussed in this paper have been implemented in commercial software for latent variable modeling. Program inputs and data for the examples discussed in this article are provided in the electronic appendix (`http://`) [included as zip file in blinded version].

## 1. INTRODUCTION

The logic of measurement invariance testing is that groups may safely be compared when their measurement parameters are exactly equal: in that case, conclusions of substantive interest are uncontaminated by measurement differences. While this logic is sound, it does not follow that comparison is never warranted when measurement parameters are not exactly equal. Moreover, in practical applications exact equality is not often plausible. These observations have led to the concept of "approximate" measurement invariance, in which a certain, "small" amount of non-invariance is permitted.

However, while these methods ensure that measurement differences are “small”, even “small” measurement differences could, in principle, still contaminate the conclusions of interest. Conversely, those measurement differences allowed for may not necessarily affect the conclusions of interest.

(Muthén and Asparouhov, 2012) (Van De Schoot et al., 2013)

(Muthén and Asparouhov, 2014)

## 2. EPC-INTEREST

## 3. DATA

[GUY: Een mooi verhaal over Inglehart values]

[Esp. Explain why it is of interest to compare the countries on GDP/capita and % women in parliament]

[WVS 2010–2012. $n = 67{,}568$ ]

[Countries: ARM AUS AZE BLR CHL CHN COL CYP DEU DZA ECU EGY ESP EST GHA IRQ JOR JPN KAZ KGZ KOR LBN LBY MAR MEX MYS NGA NLD NZL PAK PER PHL POL QAT ROU RUS RWA SGP SVN SWE TTO TUN TUR UKR URY USA UZB YEM ZWE]

Table 1 shows the ranking tasks asked in the WVS 2010–2012.

| Option # | Option wording |
|---|---|
| *Set A* | |
| 1. | A high level of economic growth |
| 2. | Making sure this country has strong defense forces |
| 3. | Seeing that people have more say about how things are done at their jobs and in their communities |
| 4. | Trying to make our cities and countryside more beautiful |
| *Set B* | |
| 1. | Maintaining order in the nation |
| 2. | Giving people more say in important government decisions |
| 3. | Fighting rising prices |
| 4. | Protecting freedom of speech |
| *Set C* | |
| 1. | A stable economy |
| 2. | Progress toward a less impersonal and more humane society |
| 3. | Progress toward a society in which Ideas count more than money |
| 4. | The fight against crime |

Table 1: Options to be ranked for the three Inglehart values ranking sets in the World Values Survey 2010–2012. The wording of the question is "People sometimes talk about what the aims of this country should be for the next ten years. On this card are listed some of the goals which different people would give top priority. Would you please say which one of these you, yourself, consider the most important?"; "And which would be the next most important?".

# 4. MODEL

To examine differences in (post)materialism with respect to GDP per capita and the proportion of women in parliament, we develop a latent variable model for ranking data. The advantages of this approach are that there is no need to establish a priori rules for combining the three ranking tasks into a measure of "postmaterialism"; that it does not presuppose a fixed number of "postmaterialism" classes; that it allows for measurement error in "postmaterialism"; and that it allows for an examination of the impact of violations of measurement invariance on the conclusions of interest. There are several approaches to modeling ranking data. We adopt a sequential approach originally due to Luce (1959) and McFadden (1974). A single-level latent variable version of these models was described by Böckenholt (2002); Moors and Vermunt (2007) applied a multilevel latent class model for rankings to European Values Study data on postmaterialism. Different approaches to modeling ranking data using latent variables can be found in Maydeu-Olivares and Böckenholt (2005) and Brown and Maydeu-Olivares (2011, 2012).

The ranking model departs by describing the joint probability of a ranking on a set such as set A (see Table 1) as

$$P(A_1 = a_1, A_2 = a_2 | X = x) = \frac{\omega_{a_1 x}}{\sum_k \omega_{kx}} \frac{\omega_{a_2 x}}{\sum_{k \neq a_1} \omega_{kx}}, \tag{1}$$

where $\omega_{kx}$ is the utility ($0 < \omega_{kx} < 1$) of response option $k$ (McFadden, 1974). Crucially, the choice for second place, $A_2$, is modeled while excluding the alternative already chosen for first place, $A_1$. In other words, we model a sequential ranking process in which first place is chosen first, then second place is chosen from the remaining options. This makes the model used here different from standard models for categorical data. The utilities $\omega_{kx}$ may differ over persons due to the latent class variable $X$ (Böckenholt, 2002, pp. 171–3), which in our case will represent "(post)materialism". This latent class variable is typically chosen to have a fixed number of categories, $T$. To restrict the utilities to their allowed range and facilitate model interpretation, it is convenient instead of the utilities $\omega_{kx}$ to estimate the log-linear parameters,

$$\ln \omega_{kx} = \tau_k + \lambda_{kx}. \tag{2}$$

For identification, we set $\sum_k \tau_k = \sum_k \lambda_{kx} = \sum_x \lambda_{kx} = 0$. Using this "effects coding" (Vermunt and Magidson, 2013), the $9(T-1)$ attribute parameters $\lambda_{kx}$ can then be interpreted as the deviation from option $k$'s average log-utility in latent class $x$.

Primary interest focuses on how the probabilities of belonging to these latent (post)materialism classes differ as a function of the country-level covariates log-GDP per capita ($Z_1$) and percentage of women in parliament ($Z_2$). At the same time, it has become common practice to control for country effects using a multilevel approach. We therefore model the latent class probabilities using the multinomial logistic regression,

$$P(X = x | Z_1 = z_1, Z_2 = z_2, G = g) = \frac{\exp(\alpha_x + \gamma_{1x} z_1 + \gamma_{2x} z_2 + \beta_{gx})}{\sum_t \exp(\alpha_t + \gamma_{1t} z_1 + \gamma_{2t} z_2 + +\beta_{tg})}, \tag{3}$$

where the country-level latent variable $G$ has been introduced. The country-level effects $\beta_{tx}$ in Equation 3 could be modeled as normal, leading to a standard multilevel model. However, since in the past very differing classes of countries have been observed (Moors and Vermunt, 2007), making the normality assumption rather implausible, we instead take $G$ to be a country-level latent class variable with $S$ classes and a freely estimated distribution. This model was introduced as the "multilevel latent class regression" model by Vermunt (2003). The main parameters of substantive interest in Equation 3 are the $2(T-1)$ logistic regression coefficients $\gamma_{mx}$.

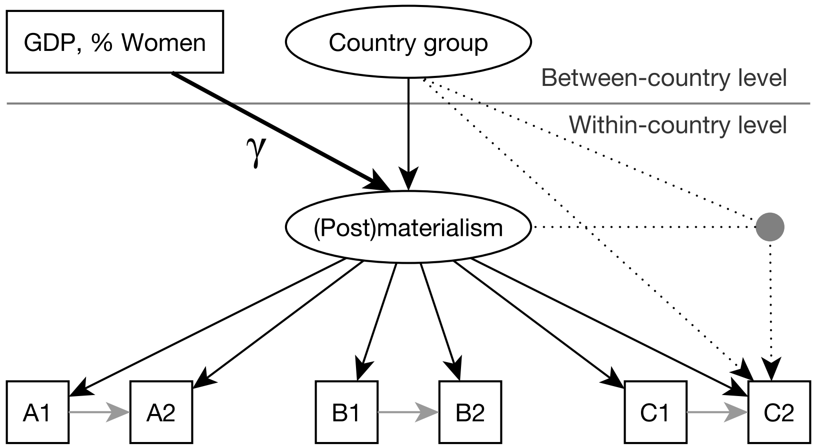


Figure 1: Graphical representation of the multilevel latent class regression model for (post)materialism measured by three partial ranking tasks. Observed variables are shown in rectangles while unobserved ("latent") variables are shown in ellipses.

Combining the three ranking tasks, each modeled according to Equation 1, with the multilevel latent class regression model in Equation 3 leads to the casewise likelihood

$$\begin{aligned} L(\boldsymbol{\theta}) &= P(A_1, A_2, B_1, B_2, C_1, C_2 | Z_1, Z_2) = \\ &\sum_G P(G) \sum_X P(X|Z_1, Z_2, G) P(A_1, A_2|X) P(B_1, B_2|X) P(C_1, C_2|X), \end{aligned} \quad (4)$$

where the realizations of variables have been dropped for clarity, and $\boldsymbol{\theta}$ is a vector collecting the $-3 + T(11 + S)$ free parameters of the model. Note that values of the within-country variables $X, A_1, A_2, B_1, B_2, C_1$, and $C_2$ may differ over persons, whereas the between-country variables $G, Z_1$, and $Z_2$ can differ only over countries.

Figure 1 shows a graphical representation of the model in Equation 4. The parameters of primary substantive interest are the effect-coded logistic regression coefficients for the effect of GDP per capita and percentage of women in parliament on the latent postmaterialism class (the stronger line labeled $\gamma$ in Figure 1). These effects of interest are controlled for the country random effects. Since these are modeled using a latent class variable, the top part of the model in Figure 1 is a multilevel latent class regression. The gray arrows between indicators within the same set are meant to show that a fixed conditional relationship is specified between the first and second choice within each set (Equation 1).

Equation 4 is, in fact, a full invariance model: it excludes potential violations of "scalar" and "metric" measurement invariance. These can be conceptualized as, respectively, a direct main effect and a direct interaction effect of the grouping variable (Mellenbergh, 1989; Kankaraš, Moors and Vermunt, 2010; Kankaraš, Vermunt and Moors, 2011). Since we do not employ multiple groups (fixed effects) to deal with country effects, but a latent class multilevel (random effects) approach, instead of considering main and interaction effects from the countries directly, we consider these for the country random effects ($G$) (see De Jong, Steenkamp and Fox, 2007; Fox and Verhagen, 2011). Thus, measurement invariance violations could be parameterized by extending Equation 2 to include country group class ($g$) effects:

$$\ln \omega_{kxg} = \tau_k + \lambda_{kx} + \tau^*_{kg} + \lambda^*_{kxg}. \tag{5}$$

The base invariance model above can then be seen as fixing the intercept deviations $\tau^*_{kg}$ and slope deviations $\lambda^*_{kxg}$ to 0. An example of this parameterization of measurement non-invariance for the second ranking of Set C ($C_2$) is shown in Figure 1 as the dotted main effect and interaction effects for scalar and metric invariance, respectively .

With three sets of three non-redundant ranking options each, there are $9(S-1)T$ possible additional $\tau^*_{kg}$ and $\lambda^*_{kxg}$ parameters representing misspecification in the full invariance model specified above. For example, with $S = 3, T = 3$, there will be 54 possible violations of invariance. This is clearly a very large number of potential violations of measurement invariance, with an astronomical ($\approx 1.8 \times 10^{16}$) number of possible subsets of non-invariant models, making the fitting of each of these possible submodels infeasible. The EPC-interest, which considers the possible effect of freeing each of the restrictions separately, is therefore an attractive alternative.

However, there are strong correlations between the possible additional parameters. First, parameters for different categories of the same variable will necessarily be strongly correlated. Second, parameters corresponding to the same ranking set (A, B, or C) will be highly correlated due to the necessary dependence between the first and second choice in a ranking task. Therefore, instead of considering only one possible misspecification at a time, we consider freeing sets of restrictions corresponding to all main effects and all interaction effects of each set. The advantages are that the estimated change in the parameters of interest will be closer to the observed change when freeing these restrictions, and that a model space to be explored of order $10^{16}$ is reduced to the examination of the effect of six sets of misspecifications on $2(T-1)$ parameters of interest, $\gamma_{jx}$.

In summary, we have attacked the problem of examining the differences in (post)-materialism over 48 countries with different levels of GDP per capita and percentage of women in parliament by formulating a multilevel latent class model for ranking data. Measurement invariance is important here because direct main and interaction effects from country groups on the ranking tasks could threaten the comparison of countries with different levels of the covariates. Therefore it becomes relevant to examine the impact of these possible violations of measurement invariance on the parameters of substantive interest using the EPC-interest. Since the data are categorical and the possible misspecifications are highly correlated, the extensions of the EPC-interest discussed in the previous section become essential.

| | | | | EPC-interest for... | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $\tau^*_{kg}$ | | | $\lambda^*_{kxg}$ | | |
| | | Estimates | | Ranking set | | | Ranking set | | |
| | | Est. | s.e. | 1 | 2 | 3 | 1 | 2 | 3 |
| Class 1 | GDP | -0.035 | (0.007) | -0.013 | 0.021 | -0.002 | **0.073** | **0.252** | 0.005 |
| Class 2 | GDP | -0.198 | (0.012) | -0.018 | -0.035 | 0.015 | -0.163 | -0.058 | 0.002 |
| | | | | | | | | | |
| Class 1 | Women | 0.013 | (0.001) | -0.006 | 0.002 | 0.000 | -0.003 | 0.029 | 0.002 |
| Class 2 | Women | -0.037 | (0.001) | 0.007 | -0.003 | 0.002 | -0.006 | -0.013 | 0.002 |

Table 2:

| | | | | EPC-interest for non-invariance of... | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $\tau^*_{kg}$ | | | $\lambda^*_{kxg}$ | | |
| | | | | Ranking set | | | Ranking set | | |
| | | Est. | s.e. | 1 | 2 | 3 | 1 | 2 | 3 |
| Class 1 | GDP | -0.127 | (0.008) | -0.015 | -0.003 | 0.002 | | | 0.097 |
| Class 2 | GDP | 0.057 | (0.011) | -0.043 | -0.013 | 0.002 | | | 0.161 |
| | | | | | | | | | |
| Class 1 | Women | 0.008 | (0.001) | -0.002 | 0.000 | 0.002 | | | 0.001 |
| Class 2 | Women | 0.020 | (0.001) | -0.007 | -0.001 | 0.002 | | | 0.007 |

Table 3:

## 5. RESULTS

Latent Gold Choice 5.0.0.14157 (Vermunt and Magidson, 2005, 2013)

LL = -418609.9616

Number of parameters (Npar) 63

Program input can be found in Appendix A, while output and data are included in the online Appendix at `http://`.

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| Class label | "Materialist" | "Postmater." | "Mixed" |
| Class size | 0.569 | 0.213 | 0.218 |
| Set A | | | |
| 1. Economic growth | 2.1102 | 0.4837 | 0.4156 |
| 2. Strong defense | -0.5285 | -1.4984 | -0.9249 |
| 3. More say | -0.5519 | 1.4683 | 0.4643 |
| 4. More beauty | -1.0298 | -0.4536 | 0.0449 |
| Set B | | | |
| 1. Order in the nation | 1.0016 | -0.5898 | 0.0435 |
| 2. More say | -0.4592 | 0.6902 | -0.2763 |
| 3. Rising prices | 0.4281 | -0.2269 | 0.3719 |
| 4. Freedom of speech | -0.9705 | 0.1266 | -0.1390 |
| Set C | | | |
| 1. Stable economy | 2.0086 | 0.0789 | 0.1715 |
| 2. Humane society | -0.7919 | 0.4450 | -0.0943 |
| 3. Ideas | -1.1402 | -0.0593 | -0.4550 |
| 4. Fight crime | -0.0765 | -0.4646 | 0.3778 |

Table 4: Attribute parameter estimates for the final model.

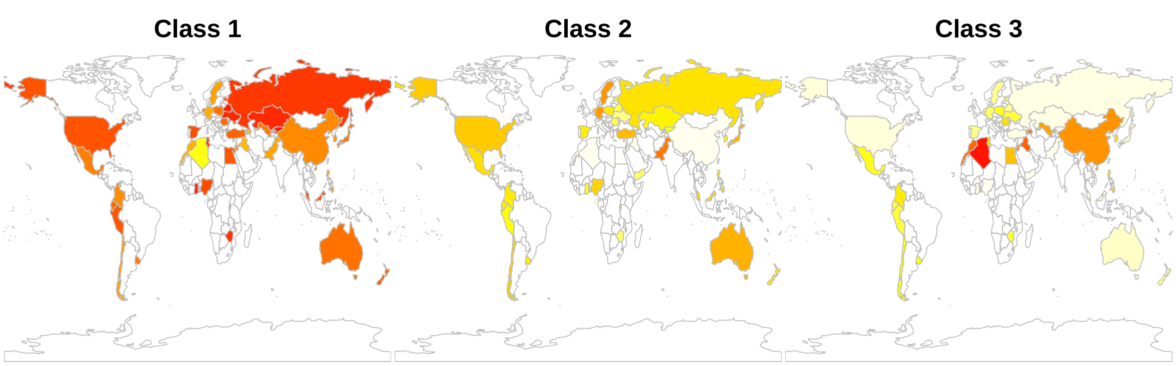


Figure 2:

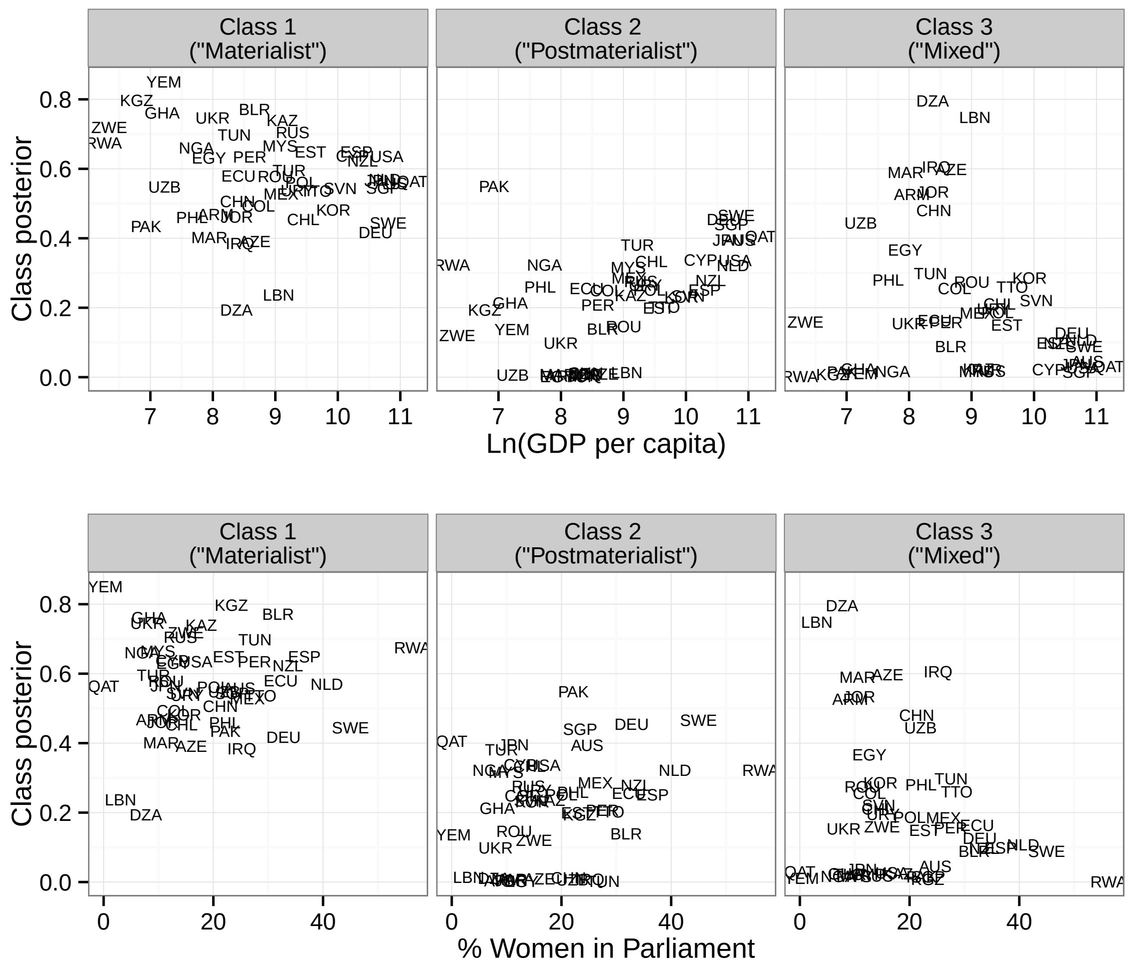
Class 1
("Materialist")
Class 2
("Postmaterialist")
Class 3
("Mixed")
Class posterior
0.0
0.2
0.4
0.6
0.8
7
8
9
10
11
Ln(GDP per capita)
Class 1
("Materialist")
Class 2
("Postmaterialist")
Class 3
("Mixed")
Class posterior
0.0
0.2
0.4
0.6
0.8
0
20
40
% Women in Parliament

Figure 3:

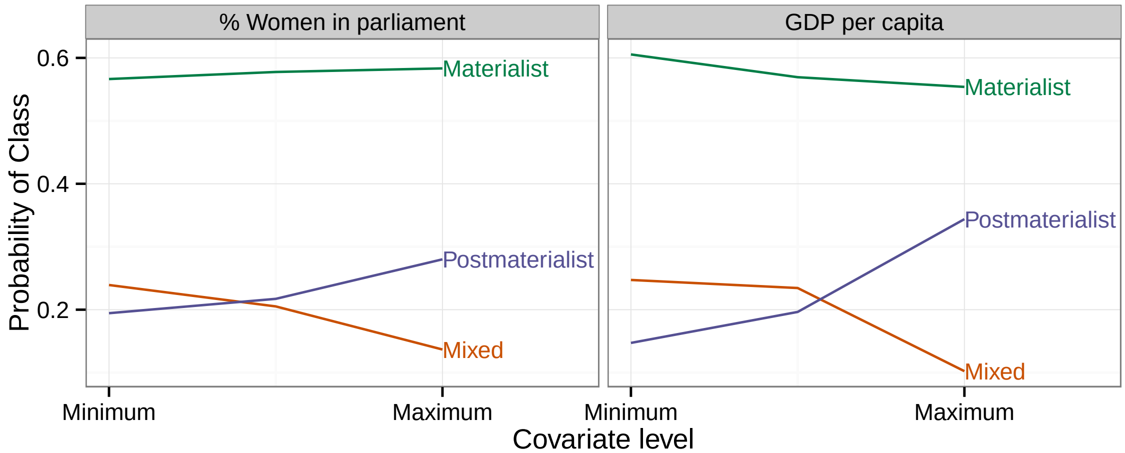


Figure 4: Estimated probability of choosing each class as a function of the covariates of interest under the final model.

## 6. DISCUSSION

## REFERENCES

Böckenholt, U. 2002. "Comparison and choice: Analyzing discrete preference data by latent class scaling models." *Applied latent class analysis*.

Brown, Anna and Alberto Maydeu-Olivares. 2011. "Item response modeling of forced-choice questionnaires." *Educational and Psychological Measurement* 71:460–502.

Brown, Anna and Alberto Maydeu-Olivares. 2012. "Fitting a Thurstonian IRT model to forced-choice data using Mplus." *Behavior research methods* 44:1135–1147.

De Jong, M.G., J.B.E.M. Steenkamp and J.P. Fox. 2007. "Relaxing Measurement Invariance in Cross-National Consumer Research Using a Hierarchical IRT Model." *Journal of consumer research* 34:260–278.

Fox, J.P. and A.J. Verhagen. 2011. "Random item effects modeling for cross-national survey data." In *Cross-cultural analysis: Methods and applications*, ed. Eldad Davidov, Peter Schmidt and Jaak Billiet. New York: Routledge pp. 467–488.

Kankaraš, Miloš, Guy Moors and Jeroen K Vermunt. 2010. "Testing for measurement invariance with latent class analysis." *Cross-cultural analysis: Methods and applications*.

Kankaraš, Miloš, Jeroen K Vermunt and Guy Moors. 2011. "Measurement equivalence of ordinal items: A comparison of factor analytic, item response theory, and latent class approaches." *Sociological Methods & Research*.

Luce, R. Duncan. 1959. *Individual Choice Behavior: a Theoretical Analysis*. New York: John Wiley and Sons.

Maydeu-Olivares, Albert and Ulf Böckenholt. 2005. "Structural equation modeling of paired-comparison and ranking data." *Psychological methods* 10:285.

McFadden, Daniel. 1974. "Conditional logit analysis of qualitative choice behavior." In *Frontiers in econometrics*, ed. Paul Zarembka. New York: Academic Press pp. 105–142.

Mellenbergh, G.J. 1989. "Item bias and item response theory." *International Journal of Educational Research* 13:127–143.

Moors, Guy and Jeroen Vermunt. 2007. "Heterogeneity in post-materialist value priorities. Evidence from a latent class discrete choice approach." *European Sociological Review* 23:631–648.

Muthén, Bengt and Tihomir Asparouhov. 2012. "Bayesian structural equation modeling: a more flexible representation of substantive theory." *Psychological methods* 17:313.

Muthén, Bengt and Tihomir Asparouhov. 2014. "IRT Studies of Many Groups: The Alignment Method." *Mplus Webnote* Version 2014-04-29.
**URL:** *http://www.statmodel.com/download/IRTAlignment.pdf*

Van De Schoot, Rens, Anouck Kluytmans, Lars Tummers, Peter Lugtig, Joop Hox and Bengt Muthén. 2013. "Facing off with Scylla and Charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance." *Frontiers in Psychology* 4:770.

Vermunt, J. K and J. Magidson. 2013. *Technical guide for Latent GOLD 5.0: Basic and advanced*. Belmont, MA: Statistical Innovations Inc.

Vermunt, Jeroen K. 2003. "Multilevel latent class models." *Sociological methodology* 33:213–239.

Vermunt, Jeroen K. and Jay Magidson. 2005. *Technical Guide for Latent GOLD Choice 4.0: Basic and Advanced*. Belmont, MA: Statistical Innovations.

## A. LATENT GOLD INPUT FOR THE FULL INVARIANCE MODEL

The input below fits the full invariance model described in the paper, setting the possible violations of invariance to zero (0). The option "score test" in the output section (only available in LG $\geq$ 5) is then used to obtain the EPC-interest values. Output and data for this example can be obtained from the online appendix at `http://`.

```
options

   maxthreads=all;
   algorithm
      tolerance=1e-008 emtolerance=0.01
      emiterations=450 nriterations=70 ;
   startvalues
      seed=0 sets=30 tolerance=1e-005 iterations=50;
   bayes
      categorical=0 variances=1 latent=0 poisson=1;
   missing  excludeall;
   output
      parameters=effect  betaopts=wl standarderrors profile
      probmeans=posterior
      frequencies bivariateresiduals estimatedvalues=regression
      predictionstatistics setprofile setprobmeans
      iterationdetails scoretest ;

choice = 3
   alternatives 'inglehart_wvs6_long.alt' quote = single
   id=alt
   choicesets 'inglehart_wvs6_long.set' quote = single
   id=set;

variables
   groupid country;
   caseid id;
   choicesetid set ;
   dependent value ranking;
   independent NY_GDP_PCAP_CD, SG_GEN_PARL_ZS;
   attribute int1 nominal, int2 nominal, int3 nominal;
   latent
      GClass  group nominal 3,
```

```
      Class nominal 3;

equations
   GClass <- 1 ;
   Class <- 1 + GClass + NY_GDP_PCAP_CD + SG_GEN_PARL_ZS;
   value <- int1 + int2 + int3 +
      int1 Class + int2 Class + int3 Class +
      (0) int1 GClass + (0) int2 GClass + (0) int3 GClass +
      (0) int1 Class  GClass +
      (0) int2 Class GClass +
      (0) int3 Class GClass ;
```