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, not all substantively relevant differences are necessarily detected by the current best practices of invariance testing. Therefore, it was recently suggested in this journal to detect the relevant set of violations of measurement invariance using sensitivity analysis. This approach 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

Ranking and paired comparisons data are central to political science, occurring for example when people rank candidates in "ranked choice" ("instant-runoff") voting systems such as the Australian federal or London mayoral elections (see e.g Karvonen, 2004; Toplak, 2010), choose from pairs of alternatives in a conjoint experiment (Hainmueller, Hopkins and Yamamoto, 2014), or rank their post-materialist value orientations in the World Values Survey (Inglehart, 1997; Inglehart and Welzel, 2010).

Such procedures yield categorical and multivariate data, and, since they preclude certain joint outcomes, lead to complex dependencies that must be accounted for.

Probabilistic models accounting for these characteristics go back at least to Thurstone (1927), who assumed choice utilities were independently normally distributed; McFadden (1974) introduced a log-linear latent utility model that relaxes this assumption. Both of these approaches treat the choices as perfectly measured; however, rankings and choices may contain measurement error, without a "true" preference being available, so that latent variable models become necessary. Maydeu-Olivares and Böckenholt (2005) extended Thurstone (1927)'s model to multivariate outcomes incorporated into a structural equation modeling framework with latent variables (see also Brown and Maydeu-Olivares, 2011, 2012); Jackson and Alwin (1980) described an earlier approach but did not account for the fact that choice data are categorical. Böckenholt (2002) extended McFadden (1974)'s model to include discrete latent variables with an arbitrary distribution – "latent class" variables.

When the results of ranking tasks are compared across countries or other groups, not only measurement error, but also "measurement equivalence" (Steenkamp and Baumgartner, 1998; Vandenberg and Lance, 2000; Schmitt and Kuljanin, 2008) becomes a concern. Inglehart's "(post)materialism" ranking tasks, for example, are often compared across countries to evaluate substantive theories on human values (e.g. Inglehart, 1997). Such country comparisons may be threatened by cross-country differences in the way respondents' ranking choices measure "(post)materialism". In other words, "measurement equivalence" means that conclusions of substantive interest should be uncontaminated by measurement differences (Oberski, 2012). To date, however, no studies have evaluated whether ranking choices are comparable across countries in the sense that parameters of interest are unaffected by possible measurement differences.

In this study, we apply a recently developed measure for this purpose, the "EPC-interest" (Oberski, 2014), to a nonparametric multilevel latent variable model for ranking data. This "multilevel latent class" (Vermunt, 2003) model for ranks or paired comparisons extends Böckenholt (2002)'s model to include multiple groups and a nonparametric random effect (Moors and Vermunt, 2007). The model is applied to 2010 World Values Survey data from 48 countries to evaluate whether latent (post)materialism classes are associated with GDP per capita and percentage of women in a country's parliament, as suggested by values theory. The "EPC-interest" measure then approximates the change in these substantive parameters of interest that would be observed if measurement equivalence restrictions were freed. It is therefore a sensitivity analysis approach to measurement equivalence (Oberski, 2014); Van De Schoot et al. (2013, p. 12) argued that it can be seen as providing a specific definition of "approximate measurement invariance" (see Muthén and Asparouhov, 2012, 2014, for other approaches). Since the EPC-interest was originally developed for the case of equality restrictions in linear structural equation models with continuous data, we extend the EPC-interest measure here to include categorical observed and latent variables.

The contributions of this paper are threefold. First, contrary to existing uses of these data, we allow for measurement error in the (post)materialism ranking tasks. Moreover, no assumptions about the distribution of these errors or the underlying scale are necessary. Second, we evaluate whether the 48 countries in the 2010 World Values Survey can be compared for the purpose of looking at the relationship between (post)materialism and correlates central to Inglehart (1997)'s theory. Third, we develop the EPC-interest measure for latent class analysis, i.e. for models with discrete observed and latent variables. This development allows for more general application of the EPC-interest approach to measurement equivalence evaluation, for example in IRT analysis (see also Oberski and Vermunt, 2013), paired comparisons data, or ordinal factor analysis.

[The next section..]

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

Table 1: Values to be ranked for the three WVS 2010–2012 ranking sets. "Materialist" concerns are marked "M" while "postmaterialist" concerns are marked "P". The wording 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?".

2. DATA AND BACKGROUND

The data used here came from the 2010–2012 World Values Survey[1] (WVS) and included $n = 67{,}568$ respondents in 48 different countries (Appendix B provides a full list of countries). The WVS questionnaire includes Inglehart (1977)'s extended postmaterialism questions, developed to measure political values priorities. This extended version includes three sets of four values (Table 1). Of these, set B in Table 1 is known as the "short scale" that is commonly used in research on values priorities. Two of the values in each set are intended to measure "materialist" concerns while the two

[1] `http://www.worldvaluessurvey.org/`

remaining values indicate “postmaterialist” political goals; for each of the three sets, WVS respondents are then asked to rank their first and second priorities.

Postmaterialism theory has been founded on the dual-hypotheses model (Inglehart, 1981, p. 881). The first hypothesis holds that an individual’s priority set reflects the socio-economic environment because one attaches relatively more importance to scarce objects. This scarcity hypothesis is complemented by a socialization hypothesis that stresses the importance of experiences in the ‘formative’ years. The latter hypothesis constitutes the heart of the argument regarding generational differences in political priorities since it indicates that, in reaching adulthood, values tend to crystallize in personality. By a process of social metabolism, society gradually changes from materialist to postmaterialist.

In line with the scarcity hypothesis, Inglehart (1997) and Inglehart and Welzel (2010) argued that socio-economic development indicators per country such as GDP per capita correlate strongly with aggregated scores of self-expression values. Moreover, in line with the socialization hypothesis, associations of aggregated values priorities with socio-cultural variations across countries such as religious heritage and gender equality have been found by Inglehart, Norris and Welzel (2002). In short, substantive interest in this literature focuses on the associations between (post)materialism and both socio-economic and socio-cultural country-level aggregates. In this study we will therefore following these authors by examining the aggregate relationship of values priorities and by using GDP per capita and the percentage of women in parliament to represent socio-economic and socio-cultural correlates of values, respectively. These country-level variables were obtained from the World Bank database[2] using the `WDI` package (Arel-Bundock, 2013) in `R 3.0.2` (R Core Team, 2012).

In their previous work, Inglehart and colleagues examined these relationships of

[2] `http://data.worldbank.org/`

interest between value priorities and country-level covariates by creating an observed categorization of the ranking choices into three groups: "materialist", "postmaterialist" and "mixed". These groups are defined *a priori*, and their usage as observed variables implies that a complete absence of measurement error is assumed. It also implies that no differences in measurement exist across the countries to be compared. In other words, any comparison of the created categories over countries is assumed to be fully due to substantive rather than measurement differences. The contribution of this study to the values priorities literature is to relax the assumption of zero measurement error, examine empirically whether three class can adequately represent the data, and apply the EPC-interest procedure to examine the implicit claim of full cross-country invariance of these measures.

3. 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 observing choice $A_1 = a_1$ on set A (see Table 1) for first place and $A_2 = a_2$ for second place 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 ω_{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 ω_{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 ω_{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 λ_{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 β_{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 nonparametric 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 γ_{mx}.

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

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

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 γ in Figure 1). These effects of interest

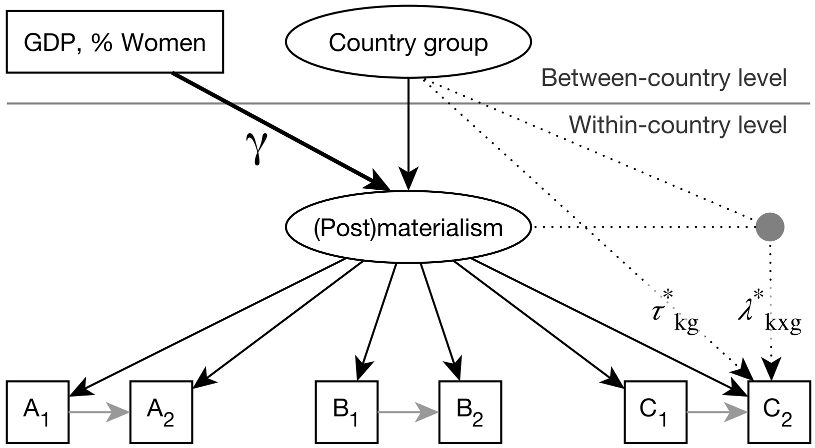

Figure 1: Graphical representation of the multilevel latent class regression model for (post)materialism measured by three partial ranking tasks. Observed variables are shown in rectangles while unobserved ("latent") variables are shown in ellipses.

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 τ^*_{kg} and slope deviations λ^*_{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 τ^*_{kg} and λ^*_{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, γ_{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. The following section will introduce a method for doing so without the need to fit the alternative models.

4. EPC-INTEREST

In the application discussed, as in any application of invariance testing, interest focuses on a particular (function of) free parameters of the model. As argued by Oberski (2014), the question of invariance testing is then relevant insofar as it affects these parameters of interest. Such parameters may differ in models that allow for non-invariance, which in the model introduced in the previous section was parameterized as direct main and interaction effects of a group random effect. Instead of investigating whether such effects are statistically significantly different from zero, we suggest to investigate the impact of such effects on the parameters of interest. This is done using the "expected parameter change" in the parameter of interest, "EPC-interest".

The "EPC-interest" estimates the change in a free parameter of the model that one can expect to observe if a particular restriction were freed. It is therefore a method of sensitivity analysis. However, the researcher is not forced to estimate all possible alternative models, but can evaluate the sensitivity of the results after fitting the restrictive full invariance model. The EPC-interest is based on the work of Saris, Satorra and Sörbom (1987), who introduced the expected parameter change in a fixed parameter for linear structural equation models (SEM), and Bentler and Chou (1993), who introduced the expected parameter change in a free parameter after freeing a fixed parameter for SEM. Oberski (2014) suggested using the EPC-interest for invariance testing using linear structural equation models with continuous data.

In this section we apply the EPC-interest to models with categorical observed and latent variables. We also extend the measure to evaluate the probable impact of several misspecifications at the same time. This last extension is particularly relevant in categorical data models, since parameters related to different categories of the same variable are often highly correlated. In such cases, the impact of misspecifications is better evaluated in sets of strongly related parameters. In the application to ranking data discussed above, for instance, parameters relating to the different ranking objects in the same set will be highly correlated, since the second choice in each ranking is modeled conditionally upon the first choice (Equation 1). The impact of freeing the parameter λ_{111} can therefore not be seen separately from that of $\lambda_{211}, \lambda_{121}$, etc. We therefore consider the impact of all possible interactions of country grouping G with the latent postmaterialism variable X jointly for each of the three ranking tasks.

The key concept is considering the likelihood not only as a function of the free parameters of the model, but also as a function of the parameters that are fixed to obtain the full invariance model. Collecting the free parameters in a vector $\boldsymbol{\theta}$ and the fixed parameters in a vector $\boldsymbol{\psi}$, we assume the likelihood can be written as an explicit function of both sets of parameters, $L(\boldsymbol{\theta}, \boldsymbol{\psi})$. The maximum-likelihood estimates $\hat{\boldsymbol{\theta}}$ of the free parameters can then be seen as obtained under the full invariance model that sets $\boldsymbol{\psi} = 0$, i.e. $\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}, \boldsymbol{\psi} = 0)$. Further, define the parameters of substantive interest as $\boldsymbol{\pi} := \mathbf{P}\boldsymbol{\theta}$, where $\mathbf{P}$ is typically a logical (0/1) selection matrix, although any linear function of the free parameters $\boldsymbol{\theta}$ may be taken. Interest then focuses on the likely value these free parameters $\boldsymbol{\pi}$ would take if the fixed $\boldsymbol{\psi}$ parameters were freed in an alternative model, $\hat{\boldsymbol{\pi}}_a = \mathbf{P} \arg\max_{\boldsymbol{\theta},\boldsymbol{\psi}} L(\boldsymbol{\theta}, \boldsymbol{\psi})$.

We now show how these changes in the parameters of interest as a consequence of freeing the fixed parameters $\boldsymbol{\psi}$ can be estimated without fitting the alternative model. Let the Hessian $\hat{\mathbf{H}}_{\mathbf{ab}}$ be the matrix of second derivatives of the likelihood with respect to vectors $\mathbf{a}$ and $\mathbf{b}$, evaluated at the maximum likelihood solution of the full invariance model, $\hat{\mathbf{H}}_{\mathbf{ab}} := (\partial^2 L/\partial\mathbf{a}\partial\mathbf{b}')|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$. The expected change in the parameters of interest is then measured by the EPC-interest,

$$\text{EPC-interest} = \hat{\boldsymbol{\pi}}_a - \hat{\boldsymbol{\pi}} = \mathbf{P}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\psi}}\mathbf{D}^{-1}\left[\left.\frac{\partial L(\boldsymbol{\theta},\boldsymbol{\psi})}{\partial\boldsymbol{\psi}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}\right] + O(\boldsymbol{\delta}'\boldsymbol{\delta}), \tag{6}$$

where $\mathbf{D} := \hat{\mathbf{H}}_{\boldsymbol{\psi}\boldsymbol{\psi}} - \hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\psi}}'\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\psi}}$ and the deviation from the true values is $\boldsymbol{\delta} := \boldsymbol{\vartheta} - \hat{\boldsymbol{\vartheta}}$, with $\boldsymbol{\vartheta}$ collecting the free and fixed parameters in a vector, $\boldsymbol{\vartheta} := (\boldsymbol{\theta}', \boldsymbol{\psi}')'$. Note that, apart from the order of approximation term $O(\boldsymbol{\delta}'\boldsymbol{\delta})$, Equation 6 contains only terms that can be calculated after fitting the invariance model. Thus, it is not necessary to fit the alternative model to obtain the EPC-interest.

In the structural equation modeling literature, the expected change in the fixed parameters $\boldsymbol{\psi}$ is commonly found and implemented in standard SEM software. This measure is commonly know as the "EPC", but to avoid confusion we term it "EPC-self" here. The EPC-self and EPC-interest both consider the impact of freeing restrictions, but differ in the target of this impact: the EPC-self evaluates the impact on the restriction itself, whereas the EPC-interest evaluates the impact on the parameters of interest. In spite of these differences, the two measures are related: this can be seen by recognizing that $-\mathbf{D}^{-1}\left[\frac{\partial L(\boldsymbol{\theta},\boldsymbol{\psi})}{\partial \boldsymbol{\psi}}\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}\right] = \text{EPC-self} \approx \boldsymbol{\psi} - \hat{\boldsymbol{\psi}}$ so that, from Equation 6,

$$\text{EPC-interest} = -\mathbf{P}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\psi}}\,\text{EPC-self} \approx -\mathbf{P}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\psi}}\left(\boldsymbol{\psi} - \hat{\boldsymbol{\psi}}\right) \tag{7}$$

Furthermore, since $\hat{\boldsymbol{\psi}}$ and $\hat{\boldsymbol{\theta}}$ are implicitly related by the fact that they are both solutions to the equation $\partial L/\partial \boldsymbol{\vartheta} = \mathbf{0}$, invoking the implicit function theorem yields $-\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\psi}} = \partial\boldsymbol{\theta}/\partial\boldsymbol{\psi}'$, so that

$$\text{EPC-interest} = \mathbf{P}\left(\frac{\partial\boldsymbol{\theta}}{\partial\boldsymbol{\psi}'}\right)\left(\boldsymbol{\psi} - \hat{\boldsymbol{\psi}}\right), \tag{8}$$

that is, the EPC-interest can be seen simply as the coefficient of a linear approximation to the relationship between the free and fixed parameters, multiplied by the change in the fixed parameters. This demonstrates the difference with the sensitivity analysis approach common in econometrics (Magnus and Vasnev, 2007, p. 168) and applied to SEM by Yuan, Marshall and Bentler (2003), in which only $\partial\boldsymbol{\theta}/\partial\boldsymbol{\psi}'$ is considered: the EPC-interest combines both the direction and the magnitude of the misspecification.

The derivation of the EPC-interest given in Equation 6 starts from the full invariance solution. We then find a hypothetical new maximum of the likelihood by setting the gradient of a Taylor expansion of the likelihood around the full invariance solution to zero:

$$\frac{\partial L(\boldsymbol{\theta},\boldsymbol{\psi})}{\partial\boldsymbol{\vartheta}} = \mathbf{0} = \begin{pmatrix} \frac{\partial L(\boldsymbol{\theta},\boldsymbol{\psi})}{\partial\boldsymbol{\theta}}\big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ \frac{\partial L(\boldsymbol{\theta},\boldsymbol{\psi})}{\partial\boldsymbol{\psi}}\big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \end{pmatrix} + \begin{pmatrix} \hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\theta}} & \hat{\mathbf{H}}_{\boldsymbol{\psi}\boldsymbol{\theta}} \\ \hat{\mathbf{H}}_{\boldsymbol{\psi}\boldsymbol{\theta}} & \hat{\mathbf{H}}_{\boldsymbol{\psi}\boldsymbol{\psi}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta} - \hat{\boldsymbol{\theta}} \\ \boldsymbol{\psi} - \hat{\boldsymbol{\psi}} \end{pmatrix} + O(\boldsymbol{\delta}'\boldsymbol{\delta}). \tag{9}$$

A similar device was used to derive the so-called “modification index” or “score test” for the significance of the hypothesis $\boldsymbol{\psi} = \mathbf{0}$ by Sörbom (1989, p. 373). Equation 6 follows directly by noting that $(\partial L(\boldsymbol{\theta},\boldsymbol{\psi})/\partial\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \mathbf{0}$ and applying the standard linear algebra result on the inverse of a partitioned matrix $\left(\hat{\mathbf{H}}^{-1}\right)_{\boldsymbol{\theta}\boldsymbol{\psi}} = -\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\hat{\mathbf{H}}_{\boldsymbol{\theta}\boldsymbol{\psi}}\mathbf{D}^{-1}$ (e.g. Magnus and Neudecker, 2007, p. 12).

The accuracy of the approximation of the EPC-interest as a measure of the change in the parameters of interest is reflected in the order of approximation term, $O(\boldsymbol{\delta}'\boldsymbol{\delta})$. It can be seen that this accuracy is quadratic in the overall change in parameters, so that the approximation can be expected to work best when the misspecifications are not "too large". This result corresponds to results on the score test ("modification index") and "EPC-self" in the literature on structural equation modeling, which can be shown to be exact under a "sequence of local alternatives", i.e. when $\boldsymbol{\vartheta} = \lim_{n\to\infty} \hat{\boldsymbol{\vartheta}} + n^{-\frac{1}{2}}\boldsymbol{\delta}$ (Satorra, 1989, p. 135). It is important to note here that $\boldsymbol{\delta}$ is the deviation from the "true" value of $\boldsymbol{\vartheta}$, rather than the deviation from the limit of the parameter estimates under the alternative model. Therefore another view on the accuracy is that it will be better when the alternative model is not strongly misspecified. For this reason it is also important to consider freeing sets of very strongly related parameters simultaneously, since a change in one of them will then imply a change in the others, and, consequently, a misspecified alternative model.

4.1. *EPC-interest for the multilevel latent class model*

The previous section discussed the rationale and derivation of the EPC-interest for general maximum likelihood problems. For the application involving categorical latent and observed variables, $\boldsymbol{\psi}$ corresponds to the direct effects τ^*_{jkg} and λ^*_{jkxg} in Equation 5 (see also Figure 1). The Jacobian of the individual likelihood with respect to the τ^*_{jkg}

parameters is then

$$\frac{\partial L_i(\boldsymbol{\theta})}{\partial \tau^*_{jkg}} = L_i(\boldsymbol{\theta}) \sum_G \left[P(G) \sum_X P(X|Z_1, Z_2)(Y_{jk}G_g - E[Y_{jk}G_g|X]) \right], \qquad (10)$$

Similarly,

$$\frac{\partial L_i(\boldsymbol{\theta})}{\partial \lambda^*_{jkxg}} = L_i(\boldsymbol{\theta}) \sum_G \left[P(G) \sum_X P(X|Z_1, Z_2)(Y_{jk}X_xG_g - E[Y_{jk}X_xG_g|X]) \right], \qquad (11)$$

where X_x is the effect-coded value for class x of the within-country latent postmaterialism variable X. Consistent estimates of all derivatives can be obtained by replacing the parameters of the model by their maximum likelihood estimates, providing the vector $[\partial L(\boldsymbol{\theta}, \boldsymbol{\psi})/\partial \boldsymbol{\psi}]|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ necessary for calculating Equation 6.

The Hessian matrix $\mathbf{H}$ The derivatives for other parameters of the model are given by Vermunt and Magidson (2013). These have a form similar to Equations 10 and 11.

[]

5. RESULTS

We now estimate the multilevel latent class model using 48 countries from the World Values Survey[3] for 2010–2012. Moors and Vermunt (2007) applied a similar model to 19 countries in the 1990 European Values Study data and investigated the relationship of the latent postmaterialism classes with political attitudes scales. However, they did not consider the possible impact of violations of measurement invariance on such estimates. Since our goal is to compare countries with different levels of GDP per capita and proportion of women in parliament, in this study we evaluate the sensitivity of

[3] http://www.worldvaluessurvey.org

Postmaterialism (X) classes, $\|\{G\}\| = 1$				Country group (G) classes, $\|\{X\}\| = 3$			
#Classes	Log-lik	#Par	BIC(L^2)	#Classes	Log-lik	#Par	BIC
1	-460512.7	9	-10346.1	1	-447646.2	29	895616.3
2	-449836.9	19	-31586.2	2	-444754.8	32	889867.0
3	-447646.2	29	-35855.8	3	-443216.6	35	886824.1
4	-446211.1	39	-38614.4	4	-442734.2	38	885892.8
5	-445246.3	49	-40432.4	5	-442436.5	41	885330.9
6	-444776.4	59	-41260.3	6	-442110.5	44	884712.3
7	-444384.0	69	-41933.4	7	-441946.2	47	884417.3

Table 2: Log-likelihood, number of parameters, and Bayesian Information Criterion (BIC) for models with different numbers of classes for the postmaterialism (within-country) and country group (between-country) latent class variables.

these effect estimates of interest to measurement invariance assumptions using the EPC-interest.

The model elaborated in the previous sections as well as EPC-interest measures for it are implemented by one of the authors in the software Latent Gold Choice version 5.0.0.14157 (Vermunt and Magidson, 2005, 2013, pp. 135–136). Program input for this analysis can be found in Appendix A, while output and data are included in the online Appendix at `http://[BLINDED]`.

In choosing the number of classes for the postmaterialism (within-country) and country group (between-country) latent class variables, we follow the advice of Lukočienė, Varriale and Vermunt (2010) to first fix the number of country-group classes to unity and choose a number of within-country classes, subsequently fixing the number of within-country classes to this chosen number and determining the number of country-group (between-country) classes. The left-hand side of Table 2 shows the log-likelihoods, number of parameters and Bayesian Information Criterion (BIC) values based on the L^2 for the model with one country-group class and an increasing number of postmaterialism classes. It can be seen that the BIC, which penalizes model complexity, decreases with each additional latent postmaterialism class. In fact, the BIC does not stop decreasing even when incrementing the number of classes to 14 (not shown in Table 2 for brevity).

In the literature on postmaterialism (e.g. Inglehart, 1997; Inglehart, Norris and Welzel, 2002; Inglehart and Welzel, 2010), the number of postmaterialism classes is typically fixed to three: "postmaterialist", "materialist", and "mixed". Clearly, using the WVS ranking tasks and imposing full invariance, many more qualitative postmaterialism classes can be distinguished than the traditional three classes. This corresponds to findings by Moors and Vermunt (2007); however, these authors (following Hagenaars, 1990) also argued that "one can safely interpret the results (...) if adding another class does not result in important changes of the latent class weights for the other classes" (p. 637). While this is a somewhat subjective criterion, the three-class solution found in the data does correspond to the theoretical "postmaterialist", "materialist", and "mixed" classes, whose parameters appear to change little in the models with a greater number of classes. Moreover, the greatest reduction in BIC seen in Table 2 takes place when moving from a one-class to a two-class model, with relatively small improvements after three or more classes. We therefore follow the theoretical literature in selecting the three-class model.

While selecting the number of country-group classes, we find that the BIC improves little after three classes (right-hand side of Table 2), so that our initial full invariance model has three postmaterialism (within-country) and three country group (between-country) classes. However, this full invariance model still imposes $\tau^*_{jkg} = 0$ and $\lambda^*_{jkxg} = 0$, i.e. there is no difference in measurement over the group random effect variable. Such violations of measurement invariance are relevant in the present analysis because they could potentially threaten the conclusions of substantive interest, namely the relationships γ between postmaterialism on the one hand and GDP per capita and percentage of women in parliament on the other.

		EPC-interest for...							
				τ^*_{jkg}			λ^*_{jkxg}		
		Estimates		Ranking task			Ranking task		
		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 3: Full invariance multilevel latent class model: parameter estimates of interest with standard errors (columns 3 and 4), as well as expected change in these parameters measured by the EPC-interest when freeing each of six sets of possible misspecifications (columns 4–10).

We therefore calculated the EPC-interest for these parameters, of which our three-class model has four: two for each for each of the two independent variables. Measurement invariance violations can potentially take the form of 6 direct main effects (τ^*_{jkg}) –$(4-1)$ category effects in $(3-1)$ classes– and 12 direct interaction effects (λ^*_{jkxg}) for each of the three ranking tasks, totaling 54 possible misspecifications in the full invariance model. However, these misspecifications are strongly correlated and should not be considered separately. Rather, we consider the probable impact on the γ parameters of interest of freeing the direct main effects for each ranking task separately and of freeing the direct interactions for each ranking task separately. In short, rather than consider the direct and interaction effects for each of the 48 countries on each of the 3 unique categories of each of the three ranking tasks, making for $4 \times 3 \times (4-1) \times (48-1) \times (1+(3-1)) = 5076$ potential EPC-interest values, we evaluate direct effects of the country group random effect and consider their impact jointly for strongly correlated misspecifications, reducing the problem to 24 EPC-interest values of interest.

Table 3 shows these 34 EPC-interest values calculated from the full invariance model. These EPC-interest values estimate the likely change from the maximum likelihood estimates (column 3 in Table 3) in the four parameters of substantive interest after freeing the direct main or interaction effects for each of the three ranking tasks (columns 5–10 in Table 3). In the full invariance model, Class 1 corresponds to a "post-materialism" class. The estimate -0.035 (s.e. 0.007) shown in Table 3) would therefore suggest that more prosperous nations tend to be less postmaterialist. This directly contradicts the theory of Inglehart (1997), which suggests that this coefficient should be positive.

Since the theory specifies only that certain coefficients should be positive or negative, the key focus of substantive interest should be whether misspecifications in the invariance model can potentially change the sign of a parameter of interest. In Table 3, we therefore look for EPC-interest values that, when added to the estimates in column 3, would change the sign of those estimates. It can be seen in the Table that two such EPC-interest are indeed present, namely the direct interaction effect of the country group class with the postmaterialism class on ranking tasks 1 and 2. The analogous finding (Kankaraš, Vermunt and Moors, 2011) if CFA had been used would be that of "metric equivalence" for ranking task 3 but not ranking tasks 1 and 2. This means that the attribute parameters that define the classes for these two tasks differ over country groups, and that after accounting for these differences the effect of GDP on postmaterialism is estimated to be positive rather than negative. This set of misspecifications is thus of substantive interest and should be amended in the model.

We therefore free these two sets of measurement invariance violations, allowing for differences in the parameters of ranking tasks 1 and 2 across country groups. Table 4 shows the substantive parameter estimates and EPC-interest interest values for the resulting partial invariance model, which has 63 parameters and a log-likelihood of -418610.0 (BIC = 837920.5). The likelihood ratio test of improvement in model fit is highly significant ($\chi^2_{\text{df}=28} = 24607$). Moreover, the substantive logistic regression

				EPC-interest for non-invariance of...					
					τ^*_{kg}			λ^*_{kxg}	
				Ranking task			Ranking task		
		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 4: Partially invariant multilevel latent class model: parameter estimates of interest with standard errors (columns 3 and 4), as well as expected change in these parameters measured by the EPC-interest when freeing each of four sets of remaining possible misspecifications (columns 4–6 and 10).

coefficient for the effect of GDP on the postmaterialism class (Class 2 in Table 4), is indeed positive after freeing the detected misspecifications. Re-calculating EPC-interest values for the remaining possible misspecifications reveals that none of the possible misspecifications in this partial invariance model has the potential to change the substantive conclusions. In this sense, we therefore conclude that the partial invariance model fits "approximately", since none of the substantive conclusions based on it are threatened by measurement invariance violations.

The measurement parameters for the final model are shown in Table 5. This table shows the sizes of the three postmaterialism classes (third row) and the "attribute parameters" λ_{jkx}, i.e. each class's difference from the average log-utility. It can be seen that for respondents in the first class, which is also the largest, the log-utility for the items "economic growth", "order in the nation", and "stable economy" is much larger than average. This class therefore clearly represents "materialists" who value economic considerations. The second class represents those for whom "more say" (set A), "more say" (set B), "freedom of speech", and "humane society" has a higher log-utility than average. This second class constituting about 21% of the respondents is therefore labeled "post materialist". Finally, preferences in the third class are mostly

	Class 1	Class 2	Class 3
Class label	“Materialist”	“Postmater.”	“Mixed”
Class size	0.569	0.213	0.218
(s.e.)	(0.0114)	(0.0179)	(0.0280)
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 5: Attribute parameter estimates for the final model.

toward the “materialist” items of sets B and C, but this class also includes people for whom “more say” is important when compared with a “strong defense” rather than when compared with “rising prices”. Therefore we apply the label “Mixed” to this class.

This combination of classes corresponds to the labels and groups created by Inglehart (1997). However, a key difference with the predefined groups of Inglehart (1997) and the latent classes in the current model is that the latent classes here arise purely from the model and the response patterns observed in the data, rather than being imposed a priori. Crucially also, our model allows for measurement error in the ranking tasks.

Finally, Figure 2 summarizes the loglinear parameter estimates of interest shown column 3 of Table 4. These graphs plot the covariate level on the horizontal axes and

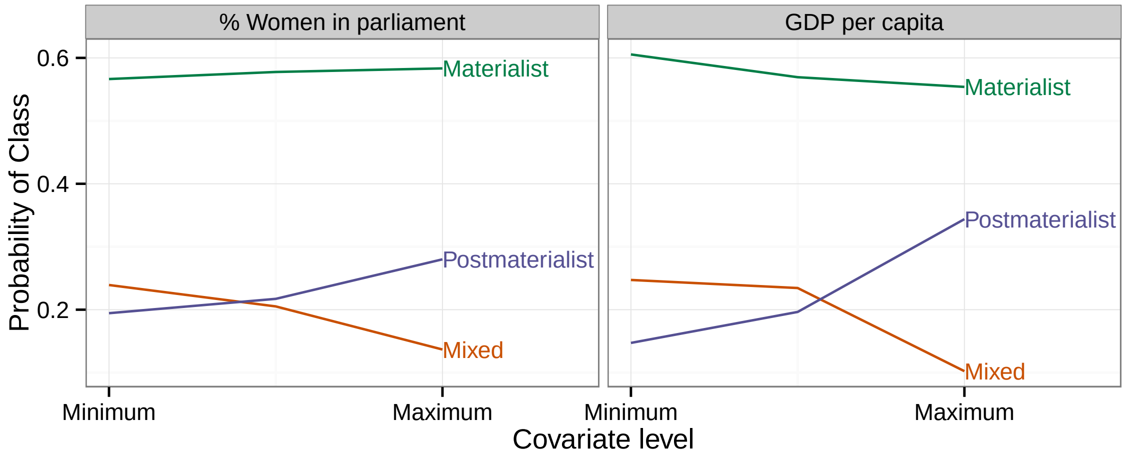

Figure 2: Estimated probability of choosing each class as a function of the covariates of interest under the final model.

the probability of belonging to each of the classes defined in Table 5 on the vertical axes. The left-hand plot in Figure 2 shows the estimated marginal relationship between the percentage of women in parliament and the postmaterialism classes. It can be seen that the probability of the "post materialist" class is much larger for countries with a large percentage of women in parliament than it is for countries for fewer female parliamentarians: the predicted probability of belonging to the postmaterialist class for countries with the observed maximum percentage of women in parliament is 0.30, which is 7 percentage points above the average "postmaterialist" class membership of 0.23 (shown in Table 5). The probability of belonging to the purely economical "materialist" class also increases very slightly, to the overall detriment of membership of the "mixed" class in which "order in the nation" and "fighting crime" is sometimes endorsed.

The right-hand side of Figure 2 plots the estimated effect of log-GDP per capita on the three class membership probabilities. A similar pattern to that for percentage of women in parliament is observed for the "postmaterialist" and "mixed" classes, while the probability of belonging to the "materialist" class decreases somewhat with GDP per capita. The differences in predicted class membership are even larger over the observed GDP range than for the percentage of women in parliament: the predicted postmaterialist class membership at the maximum observed GDP is 0.40, which is 17 percentage points above the average. Greater prosperity is thus clearly strongly associated with postmaterialism.

6. DISCUSSION

In summary, the present application of a multilevel latent class model to the WVS postmaterialism ranking tasks has shown how the EPC-interest may be used to draw stable substantive conclusions when measurement invariance is a concern. Especially encouraging is the fact that the full invariance model provides estimates that contradict Inglehart and Welzel (2010)'s values theory, whereas applying the procedure suggested here happens to yield estimates that corroborate it. This is not because we have adjusted the model to fit the theory but rather because freeing those misspecifications that could invalidate the conclusion leads to this result. The EPC-interest is a useful guide for this purpose, reducing a potentially 5076-dimensional problem to the examination of Tables 3 and 4.

REFERENCES

Arel-Bundock, Vincent. 2013. *WDI: World Development Indicators (World Bank)*. R package version 2.4.

URL: *http://CRAN.R-project.org/package=WDI*

Bentler, P.M. and C.P. Chou. 1993. "Some new covariance structure model improvement statistics." In *Testing structural equation models*, ed. K.G. Jöreskog and J. Scott Long. Sage focus editions Sage Publications pp. 235–235.

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.

Hagenaars, Jacques A P. 1990. *Categorical longitudinal data: Log-linear panel, trend, and cohort analysis*. Newbury Park: Sage.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: understanding multidimensional choices via stated preference experiments." *Political Analysis* 22:1–30.

Inglehart, Ronald. 1977. *The silent revolution*. Princeton: Princeton University Press.

Inglehart, Ronald. 1981. "Post-materialism in an environment of insecurity." *The American Political Science Review* 75:880–900.

Inglehart, Ronald. 1997. *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*. Cambridge, UK: Cambridge University Press.

Inglehart, Ronald and Christian Welzel. 2010. "Changing mass priorities: The link between modernization and democracy." *Perspectives on Politics* 8:551–567.

Inglehart, Ronald, Pippa Norris and Christian Welzel. 2002. "Gender equality and democracy." *Comparative Sociology* 1:321–345.

Jackson, David J and Duane F Alwin. 1980. "The factor analysis of ipsative measures." *Sociological Methods & Research* 9:218–238.

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*.

Karvonen, Lauri. 2004. "Preferential voting: incidence and effects." *International Political Science Review* 25:203–226.

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

Lukočienė, Olga, Roberta Varriale and Jeroen K Vermunt. 2010. "The simultaneous decision(s) about the number of lower-and higher-level classes in multilevel latent class analysis." *Sociological Methodology* 40:247–283.

Magnus, J.R. and A.L. Vasnev. 2007. "Local sensitivity and diagnostic tests." *The Econometrics Journal* 10:166–192.

Magnus, J.R. and H. Neudecker. 2007. *Matrix Differential Calculus with Applications in Statistics and Econometrics, Third Edition.* New York: John Wiley & 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*

Oberski, D.L. 2012. "Comparability of Survey Measurements." In *Handbook of Survey Methodology for the Social Sciences*, ed. Lior Gideon. New York: Springer-Verlag pp. 477–498.

Oberski, D.L. 2014. "Evaluating Sensitivity of Parameters of Interest to Measurement Invariance in Latent Variable Models." *Political Analysis* 22:45–60.

Oberski, D.L. and J.K. Vermunt. 2013. "A Model-Based Approach to Goodness-of-Fit Evaluation in Item Response Theory." *Measurement: Interdisciplinary Research & Perspectives* 11:117–122.
URL: *http://daob.nl/wp-content/uploads/2013/07/Oberski-Vermunt-A-model-based-approach-to-goodness-of-fit-evaluation-in-item-response-theory-2013-07-28.pdf*

R Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
URL: *http://www.R-project.org/*

Saris, W.E., A. Satorra and D. Sörbom. 1987. "The Detection and Correction of Specification Errors in Structural Equation Models." *Sociological Methodology* 17:105–129.

Satorra, A. 1989. "Alternative Test Criteria in Covariance Structure Analysis: A Unified Approach." *Psychometrika* 54:131–151.

Schmitt, N. and G. Kuljanin. 2008. "Measurement invariance: Review of practice and implications." *Human Resource Management Review* 18:210–222.

Sörbom, D. 1989. "Model modification." *Psychometrika* 54:371–384.

Steenkamp, JBEM and H. Baumgartner. 1998. "Assessing measurement invariance in cross-national consumer research." *Journal of Consumer Research* 25:78–107.

Thurstone, Louis L. 1927. "A law of comparative judgment." *Psychological review* 34:273.

Toplak, Jurij. 2010. Preferential voting: definition and classification. In *Annual Meeting of the Midwest Political Science Assication 67th Annual National Conference*.

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.

Vandenberg, R.J. and C.E. Lance. 2000. "A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research." *Organizational Research Methods* 3:4–70.

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.

Yuan, K.H., L.L. Marshall and P.M. Bentler. 2003. "Assessing the effect of model misspecifications on parameter estimates in structural equation models." *Sociological Methodology* 33:241–265.

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=0 latent=0 poisson=0;
   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 ;
```

B. COUNTRIES INCLUDED IN THE STUDY

ISO3 code	Country name	ISO3 code	Country name
DZA	Algeria	MAR	Morocco
ARM	Armenia	NLD	Netherlands, The
AUS	Australia	NZL	New Zealand
AZE	Azerbaijan	NGA	Nigeria
BLR	Belarus	PAK	Pakistan
CHL	Chile	PER	Peru
CHN	China	PHL	Philippines
COL	Colombia	POL	Poland
CYP	Cyprus	QAT	Qatar
ECU	Ecuador	RUS	Russian Federation
EGY	Egypt	RWA	Rwanda
EST	Estonia	SGP	Singapore
DEU	Germany	SVN	Slovenia
GHA	Ghana	ESP	Spain
IRQ	Iraq	SWE	Sweden
JPN	Japan	TTO	Trinidad and Tobago
JOR	Jordan	TUN	Tunisia
KAZ	Kazakhstan	TUR	Turkey
KOR	Korea, Republic of	UKR	Ukraine
KGZ	Kyrgyzstan	USA	United States
LBN	Lebanon	URY	Uruguay
LBY	Libya	UZB	Uzbekista
MYS	Malaysia	YEM	Yemen
MEX	Mexico	ZWE	Zimbabwe