Country Effects in ISSP-1993 Environmental Data: Comparison of SEM Approaches

Pilar Rivera
Universidad de Zaragoza

and

Albert Satorra
Universitat Pompeu Fabra. Barcelona

January, 2000

Abstract[1]

Structural equation models (SEM) are commonly used to analyze the relationship between variables some of which may be latent, such as individual "attitude" to and "behavior" concerning specific issues. A number of difficulties arise when we want to compare a large number of groups, each with large sample size, and the manifest variables are distinctly non-normally distributed. Using an specific data set, we evaluate the appropriateness of the following alternative SEM approaches: multiple group versus MIMIC models, continuous versus ordinal variables estimation methods, and normal theory versus non-normal estimation methods. The approaches are applied to the ISSP-1993 Environmental data set, with the purpose of exploring variation in the mean level of variables of "attitude" to and "behavior " concerning environmental issues and their mutual relationship across countries. Issues of both theoretical and practical relevance arise in the course of this application.

[1] This work was carried out at the ZA-EUROLAB at the Zentralarchive für Empirische Sozialforschung (ZA), Cologne. The ZA is a Large Scale Facility (LSF) funded by the Training and Mobility of Researchers (TMR) programme of the European Union. We thank Nick Longford for his comments on an earlier version of the paper. Support from Spanish grants DGES PB96-0300 and PB97-1009 is also acknowledged.

1. Introduction

Structural Equation Models (SEM) are widely used in empirical research to investigate interrelationships among variables, some of which may be latent (see, e.g., Bollen, 1989, and references contained therein). In some applications, interest focuses on the relationship between latent variables of “attitude” to and “behavior” concerning specific issues (e.g., Fishbein and Ajzen, 1975; Ajzen and Fishbein, 1982; Pieters, 1988). This relationship is of paramount importance in social environmental research where, by inducing a change in “attitude” (e.g. via education), we aim to change “behavior” and induce the individual to participate in environmental protection issues. The present paper has two purposes. A substantive aim, namely to identify variation in the mean level of variables of “attitude” to and “behavior” concerning environmental issues in a number of countries and the variation in the country level associations; and a theoretical aim, which is to compare alternative approaches to SEM analysis when we have a large number of groups, a large sample size in each group and non-normality in manifest variables. To do this, we analyzed data extracted from the ISSP-1993 survey for 22 countries detailed below.

Here, we are concerned with methodological issues such as the comparative performance of alternative methods for handling non-normal data, the validity of normal theory inferences, and the effect of large sample size on goodness of fit testing. Methodological issues also arise in selecting the model, for example when choosing between the multiple group or MIMIC model approach.

The paper is organized as follows. Section 2 describes alternative approaches and estimation methods, with a review of asymptotic robustness of normal theory methods that might be of relevance to the questions posed here. Section 3 presents the data to be analysed. Section 4 discusses the family of models considered. Section 5 describes the main results of the empirical analysis. Section 6 concludes the paper.

2. Statistical issues with non-normal data

As a standard, SEM with continuous variables is fitted by the normal theory (NT) maximum likelihood (ML) approach, or the equivalent generalized least squares method. ML analysis has been available for many years in standard computer software for SEM, such as LISREL (Jöreskog and Sörbom, 1996) and EQS (Bentler, 1995). For SEM in multiple group data, NT methods were introduced by Jöreskog (1971) and Sörbom (1974).

For non-normally distributed data, though comprising variables of a continuous scale, asymptotic distribution-free (ADF) methods have been proposed (Browne, 1984). The ADF approach is available in the above mentioned software for SEM analysis and can also be used with multiple group data. Even though the ADF approach ensures asymptotic optimality (AO) regardless of the distribution of the data, it is not very robust with small data sets. In the case of large models and a sample that is not very large, ADF suffers from computational burden such as non-convergence and/or improper solutions. For small samples, the ADF estimators are biased, as are the estimators of their standard errors. See Kaplan (1991), Finch, West and MacKinnon (1997) and Green, Akey, Fleming, Hershberger and Marquis (1997) for recent studies that evaluate alternative methods of estimating and testing in SEM, as well as a partial evaluation of the ADF approach.

An alternative to the ADF approach is pseudo maximum likelihood (PML) analysis, which uses the NT maximum likelihood estimation in conjunction with robust standard errors (se) and robust and scaled chi-square test statistics. See Satorra (1992) for an overview and discussion of PML analysis in mean and covariance structures. In contrast to the ADF approach in which fourth order sample moments are used in computing both parameter estimates and se and test statistics, in the PML approach the fourth order sample moments are used only for computing se and test statistics (parameter estimates are the same as in ML). For small samples, PML is much more robust than ADF, without substantial loss of efficiency. In combination with PML analysis, an alternative to the robust chi-square goodness-of-fit test is the scaled chi-square goodness test (Satorra and Bentler, 1994). See Satorra and Bentler (1999) for the scaled version of the difference chi-square test statistic.

An additional argument in support of the use of ML rather than ADF is provided by recent results regarding the asymptotic robustness (AR) of normal theory methods. Results for AR guarantee that, under certain conditions, normal theory produces correct inferences even when the normality assumption does not hold. See Satorra (1993) for the theory of AR in the general case of multiple group analysis. The conditions for AR are: 1) a model condition (MC), by which no restrictions are imposed on the variances and covariances of the non-normal constituents of the model; and 2) a distribution condition (DC), of mutual independence among random constituents of the model such as common and unique factors, instead of the usual assumption of no correlation. The key result about AR is that when both MC and DC hold, the NT standard errors (i.e. those of the standard ML) are consistently estimated for parameters that are not variances and covariances of non-normal constituents of the model (e.g., loadings and regression coefficients), and the NT chi-square goodness-of-fit tests (i.e., the standard log-likelihood-ratio test statistic) is appropriate. When this is the case, the robust statistics associated with PML are not required, as they coincide (asymptotically) with those of the NT. The AR theory is of particular relevance here since most of the models to be considered below verify the MC.

Up to this point, our discussion of SEM analysis has been concerned solely with continuous (CO) variables. SEM is often undertaken with categorical variables defined on ordinal scales, such as the Likert scale. Methods that take into account the OR nature of the variables are nowadays available in standard software for SEM, such as LISREL (Jöreskog and Sörbom, 1996), EQS (Bentler, 1995) and Mplus (Muthén and Muthén, 1998). Seminal studies for the analysis of SEM with OR variables include Muthén (1984), Christoffersson (1975) and Olsson (1979). Here we also compare CO and OR methods of analysis when used in the specific ISSP-1993 data set. In this context, a basic issue is whether 4 or 5 Likert-scale variables can be treated as CO variables without substantial distortion of the SEM analysis. See Coenders, Saris and Satorra (1997) for a Monte Carlo evaluation of alternative approaches to SEM with ordinal data, including a comparison of the CO and OR approaches.

When the method recognizes the OR nature of the variables, the usual Pearson correlation matrix can be replaced by a matrix that is a mix of Pearson, polychoric and polyserial correlations, according to the scale of the variables involved. Associated with this mixed correlation matrix, there is a specific matrix of asymptotic covariances. Here, we examine the approach whereby the ML fitting function is used with the mixed correlation matrix replacing the covariance matrix. The se and test statistics can then be computed by the robust analysis associated with the consistent estimate of the variance matrix of this mixed covariances. Here we carry out this analysis by combining PRELIS (to compute the mixed correlation matrix and their asymptotic covariances) and LISREL (to obtain estimates and test statistics). We call this the OR-PML since it parallels the approach of PML analysis with CO data. In fact, it uses the same formulae with appropriate substitution of the covariance matrix and the matrix of asymptotic covariances of the CO case to those corresponding to the OR case[2].

A special feature of the empirical analysis to be discussed below is its large sample size. In large samples, models are usually rejected by the chi-square goodness-of-fit test. The classic argument is that since the chi-square goodness of fit uses the sample size as a multiplying factor, any small (or substantively irrelevant) specification errors will be magnified by the large value of sample size with the result of a significant chi-square goodness-of-fit test. It has even been argued that for a very large sample size the chi-square goodness-of-fit test should be abandoned in favor of goodness-of-fit indices. However, we found that, despite a large sample size, and the fact that indeed models are only approximate, the classic chi-square goodness-of-fit test statistic plays a crucial role in model assessment (whereas typical indexes for goodness of fit were generally found to be insensitive to substantial misfits in the model). See Saris, den Ronden and Satorra, (1987) for a discussion of this issue and its relation to the power of the chi-square goodness-of-fit test.

3. Data used in the empirical analysis

The data used in the empirical analysis is extracted from the ISSP-1993 Environment file[3]. It includes variables V1, V2 and V3, indicators of “attitude” toward environmental issues (environmental conservation, willingness to sacrifice time, prices and resources, respectively), and V4 the sole indicator of “behavior”, actual participation in recycling. Variables V1 - V3 are on a 5-point Likert scale, and V4 is on a 4-point Likert scale. Appendix 1 shows the wording used in the ISSP-1993 survey for those variables. We used samples from 22 countries. The sample size for each country is shown in the first column of Table 6.

The data analyzed resulted from list-wise deletion in variables V1-V4. The extend of missing data varies from 61 % in Russia to 3% in Philippines. Overall, 25 %. of the data items are missing, most of them for variable V4, specially in countries where no recycling is available. The five highest values - Russia (61%), Poland (54%), Bulgaria (52%), Czech Republic (51 %) and Slovenia (43 %) – are recorded in countries with high values of missing data on variable V4 as a result of the response category “recycling not available” (see Q. 19a in Appendix 1). In fact, if we consider all 22 countries, 75% of missing data was due to that response. Our approach in the empirical analysis was to ignore the missing mechanism and to treat the available data as if they were complete[4].

Inspection of the marginal distributions of variables V1 to V4, overall and across countries, showed that, even though there was a clear non-normality, no extreme skewnes or ceiling

---

[2] For the sake of comparison, Table 3 shows the chi-square goodness-of-fit test provided by the WLS of LISREL with categorical data.

[3] The data is available from the Zentral Archive at the University of Cologne, Bachemer Str. 40, D-50931 Koln, Germany.

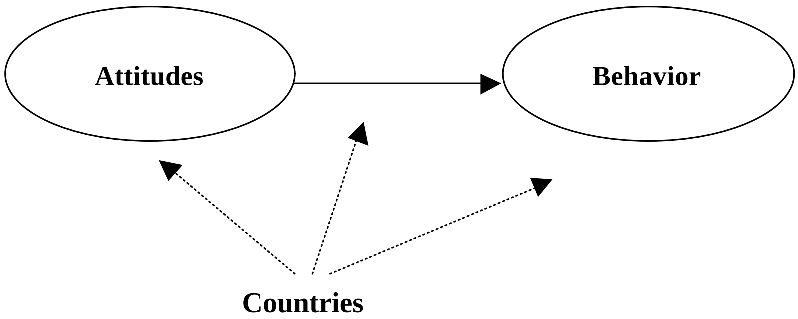
[4] The robustness of our results to this specific missing data problem awaits further study.

effects were encountered. The variables V1-V4 are ordinal (OR), with four or five categories, so it is tempting to treat them as CO variables. Thus, the question of the comparison of CO versus OR types of analysis arises. In the case of CO variable methods, the AR theory raises the additional question of whether NT methods give correct results despite the clear non-normality of the data.

4. Models

The aim of the paper is to assess country effect to both a) (mean) level of "attitude" and "behavior" and b) effect of "attitude" on "behavior".

Figure 1. Aims

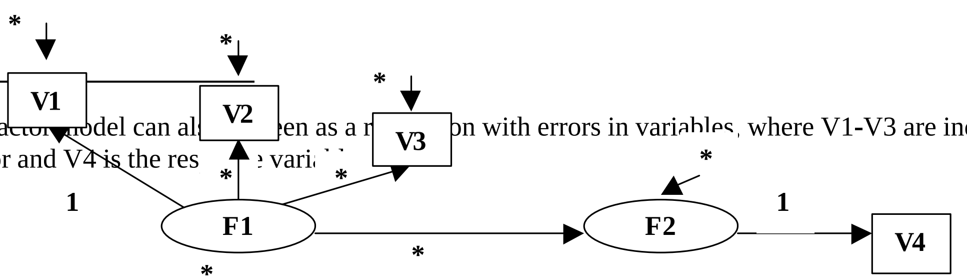


The basic structure underlying all the models considered here is shown in Figure 2. It comprises a regression equation in which factors F1 (Attitudes) and F2 (Behavior) are respectively the explanatory and dependent variables. We then have measurement equations where variables V1-V3 are indicators of F1, while V4 is a single indicator for F2. Table 1 shows the structural and measurement equations in format that is more suitable for interaction with typical software for SEM analysis.

Note that the asterisks in Table 1 parallel those in Figure 2 and represent the model's free parameters. When covariances alone are analyzed, we put zeros in the columns of the variable intercept (INT).

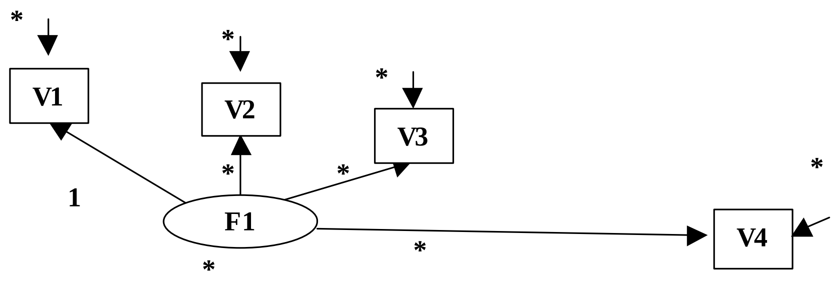
The above two-factor model is actually a single-factor model with four indicators, where the fourth loading (that of variable V4) corresponds to the regression coefficient in Figure 2. The single-factor model is depicted in Figure 3. From now on, we shall use only this single-factor model formulation, on the understanding that loadings include the regression coefficient. To ensure identification of the factor scale, the loading of variable V1 is set equal to 1[5].

Figure 2. Basic Model



[5] This one factor model can also be seen as a regression with errors in variables, where V1-V3 are indicators of the regressor and V4 is the response variable.

Figure 3. The Single-Factor Model



In order to characterize the variation across countries of the mean of factor and/or manifest variables, we need to include the means of manifest variables within the analysis. The models that are considered here are listed in Table 2 and discussed below.

4.1 Multiple group models

Various multiple group models arise depending on whether we impose or not restrictions on the means of manifest and latent variables, and the degree of invariance across groups of all model parameters.

We shall first consider the case of models with no restrictions on the means, i.e. the case of covariance structures (means excluded from the analysis). This leads to multiple group (MG) single-factor models with varying degrees of parameter invariance across groups: MG0, all parameters invariant across groups; MG1, invariance of loadings only (i.e., variances of factor and errors are set free across groups); MG2, all parameters free across groups; and finally, MG12 , invariance across groups of a selected subset of loadings. We also consider a model MG2-bis that allows for correlated errors in some countries (the countries that in the last column of Table 7 show a number different than 0). Note the nested sequence: MG0 < MG1 < MG12 < MG2, where "A < B" denotes model "A" is nested within model "B" (i.e., model "A" is model "B" with some restrictions on the parameters)

When we allow restrictions on the means, i.e., when we specify mean and covariance structures, we obtain the nested sequence of mean restricted multiple group (MMG) models: MMG0 < MMG1 < MMG12. Depending on the level of invariance across groups of the mean intercept parameters, various models arise. Here, MMG1 denotes the mean restricted multiple group model with loadings invariant across groups and variances of factors and errors free to vary across groups; MMG0 is the model with all the loadings invariant across groups; MMG12 denotes the model with only a subset of loadings invariant across groups. Restrictions on means implies adding intercept parameters (asterisks) in all the cells of the columns INT of Table 1 **(**except for cell INT -F2, as this cell is redundant with INT - V4). Models MMG0 and MMG1 impose invariance across groups of all the intercept parameters except for INT - F1 (country effect on F1) which is allowed to vary for each country. This parameter INT - F1 is the mean level of factor F1 relative to a baseline model. We choose Australia as the baseline, so all its location parameters are set to zero. Model MMG12 allows variation across group of a selected set of effects of INT to manifest variables, in addition to INT-F1. Note that the large number of groups (22) gives rise to mean structures with considerable variation on degrees of freedom. See Kaplan and George (1995) for a discussion of power issues when testing across group invariance of mean levels of the factor.

In relation to the AR theory discussed above, we note that the MC (model condition) holds for models MG1, MG12 and MG2, and the corresponding MMG models, since no restrictions are placed on variances and covariances of factors and errors, not even across groups. Thus, for these models, when DC holds (i.e. factors and errors are mutually independent, not only un-correlated), the NT analysis gives valid se for loadings and intercept parameters (including the country effects), and valid chi-square goodness-of-fit tests. Note, however, that MC does not hold for models MG0 or MMG0, as the latter restrict variances of factors and errors so that they must be equal across groups. Thus, for models MG0 and MMG0, non-normality of factor and errors may distort the validity of NT se of all parameters and also the chi-square statistics (i.e., robust se and test statistics of PML analysis are required for these models).

4.2 The MIMIC approach

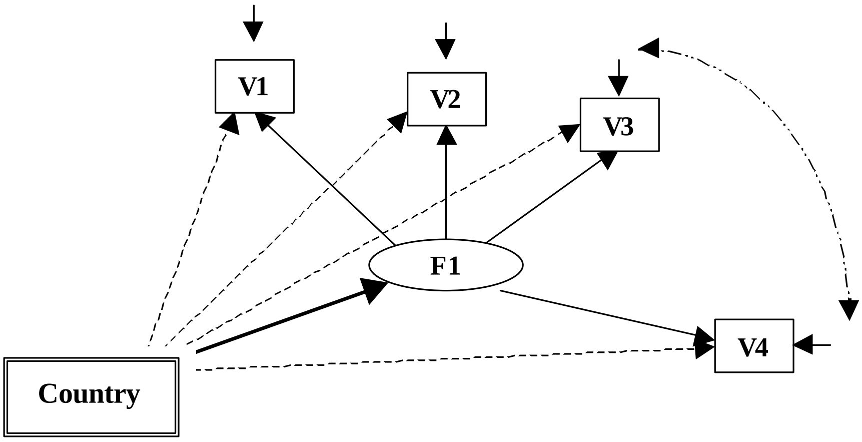
For modeling heterogeneity across groups, an alternative to multiple group analysis is the multiple indicator multiple cause (MIMIC) model approach as proposed in Muthén (1989). See Jöreskog and Goldberger (1975) for a seminal study on the MIMIC model. Muthén (1989) argued that the MIMIC model approach provides a convenient way to assess heterogeneity across groups when we have a large number of groups.

The family of MIMIC models considered here is depicted in Figure 4, where we have a single-factor model underlying four manifest variables, and there are country "effects" both on the factor and to the manifest variables. In this graph, the triangle labeled "country" represents a set of dummy variables, one less than the number of countries[6].

Table 2 also lists the various MIMIC models considered: MIMIC – A, the arrow from country to F1 (arrow in bold); MIMIC-B, arrows from country to F1 and to manifest variables (bold and dashed arrows); MIMIC-C, correlated error between V3 and V4 in addition to the effects from country (bold and dashed arrows, and the two-headed arrow – dots and dashes). Note also the nested structure: MIMIC-A < MIMIC-B < MIMIC - C.

[6] We adopt the convention of using a triangle to represent the set of dummies associated to a categorical variable. To facilitate implementation of this MIMIC model approach, we feel that SEM software should incorporate simple features to create dummy variables from a categorical variable.

Figure 4. MIMIC Models



Note that each arrow emanating from the triangle is in fact a set of arrows - one specifically for each country, except for the chosen baseline country. Some of these arrows may not be significant in a specific analysis, and can be set to zero. In fact, in our empirical analysis, the choice of the final set of direct effects from countries on manifest variables has been made on the bases of Lagrange Multiplier test statistics when fitting model MIMIC-A Note that the arrows from countries affect only the levels of the variables, not their variances or covariances.

It is interesting to compare the MIMIC modeling approaches with that discussed above for multiple group factor models. Clearly, a similar structure is imposed by MIMIC-A and MMG0, as in both cases loadings and variances must be equal across groups. There is however a fundamental difference between the multiple group models MMG1 or MG1 and the MIMIC model, as the former allow for variation across groups of the variance parameters. That is, the MIMIC approach assumes homogeneity across groups of loadings and variance parameters. In fact, Muthén (1989) pointed out that " the MIMIC approach is restricted to modeling under the assumption of a group-invariant covariance matrix for the observed response variables '' (p. 564).

The test of the assumption of invariance of variances and loadings can be carried out by comparison of the chi-square goodness-of-fit values of models MG1 and MG0. Note, however, that this test is not robust to the violation of the normality assumption, since it implies testing the equality of variances of non-normal constituents of the model. In such a case, the scaled difference of chi-square statistics as proposed recently by Satorra and Bentler (1999) could be useful.

Here we conjecture the usefulness of the MIMIC model approach beyond that restriction of equality across groups of variances of common and unique factors. Indeed, models MIMIC-A and MMG1 differ in the sense that the latter allows for heterogeneous variances across groups. We conjecture that as regards to inferences on loadings, the results of MIMIC-A are accurate approximations of those obtained by MMG1. Such robustness of the MIMIC approach to the violation of the homogeneity assumption of variances is of practical relevance, since, in comparison with the multiple group approach, the MIMIC approach is easier to implement and gives results that are more easily understandable. Note, however, that we are not suppressing the assumption of invariance of loadings across groups. Our claim is that inferences on loading parameters and effects from independent variables provided by the MIMIC approach are valid, for all practical purposes, provided it holds the assumption of an "strongly factorial invariant model" in the sense of Meredith (1993).

Finally, we should point out that the above MIMIC models satisfy the MC for AR, thus when common and unique factors are mutually independent (not only un-correlated), the NT analysis provides valid inferences for estimates of loadings and effects from independent variables[7]. The usual NT chi-square should also be valid in this case.

5. Empirical results

In this section we describe the main results of the empirical analysis when applying the models described above. For the results corresponding to CO analysis, we used EQS (Bentler, 1995), version 8, and for the OR type of results we used PRELIS and LISREL (Jöreskog and Sörbom, 1996), version 10.

We begin by reviewing the results of model assessment. Table 3 gives a summary of the chi-square goodness-of-fit values of the models listed in Table 2. As we are interested in the comparative analysis of alternative methods of estimation, we list the various alternative chi-square test statistics. Note the nested structure by rows, thus the difference of chi-square values across rows gives test statistics of various restrictions across groups. Clearly, the hypothesis of equality of variances of common and unique factors is rejected by the data (compare the chi-square values for MG0 – MG1, clearly significant). We find that a correlated error is needed between variables V3 and V4, present in models MG2-C and MG2-C bis. Clearly, the multiple group model MG2-C bis, with correlated errors in a selected set of countries, has a non-significant chi-square goodness-of-fit test, despite the large sample size. The hypothesis of invariance across groups of loadings is also rejected (compare models MG2- MG2-C). This lack of invariance across groups of loadings questions theoretically the validity of the MIMIC model for our data, but the comparison of estimates of the MIMIC and multiple group models allow us to conclude that the MIMIC model approach is useful despite the violation of the homogeneity hypothesis. In fact, for the single group analysis, we do not obtain a model that passes the chi-square goodness of fit test (at 5% significance level), not with MIMIC-C, nor with SG-C (irrespective of the method of estimation used); we believe that this might be due to the loadings heterogeneity across groups, visible in the fit of model MG2-C.

Application of the scaled difference chi-square statistic (Satorra and Bentler, 1999) for testing the null hypothesis of invariance across groups of variances of common and unique factors yields a chi-square value of 2449.35 which, in relation to 105 (= 212 – 107) degrees of freedom, clearly rejects the null hypothesis As argued by Satorra and Bentler (1999), the scaled difference is not necessarily equal to the difference of scaled chi-square statistics (=2582.98). Despite the rejection of this hypothesis of homogeneity of variances across groups, we see that,

[7] Note that non-normality of the common and unique factors could in fact be due to heterogeneity across groups of factor and errors variances.

with regard to loadings and country effects, the results for model MIMIC-A differ only slightly from those obtained using the multiple group model MMG1.

The values of the goodness-of-fit statistics are remarkable similar for the CO and OR approaches. In fact, for both approaches, we reach similar conclusions regarding model choice. Both approaches coincide in identifying the need for a correlated error among variables V3 and V4, in rejecting the hypothesis of homogeneity across groups of variances of common and unique factors, as well as in rejecting the hypothesis of homogeneity of loadings across groups.

An overview of parameter estimates for the different models and estimation methods is provided in Table 4. We see the small discrepancy in loading estimates when we use different models, and also the small discrepancy in NT and robust se, especially for the loading parameters in the models where AR results are applicable. Some differences are observed across groups in the estimated variances of unique and common factors[8], while the estimated common loadings arising from MMG1 and MMG0 are close to each other and those estimated using the single group approach of model MIMIC-A. The similarity of values applies not only to parameter estimates but also to the NT se. This is of importance in practice, since it shows that when estimating loadings and regression effects, the analysis provided by MIMIC-A is robust to the customary assumption of invariance across groups of all the parameters of the covariance structure.

In Table 5 we show the standardized values of the loadings for different models and estimation methods. Estimates of loadings (standardized values) when we use CO are generally smaller than the estimates when we use the OR approach. The difference is very small in the case of the loading for V4 (regression parameter). In any case, the differences between CO and OR estimates can be considered small in terms of their substantive relevance.

Table 6 gives the results for the multiple group models with the lowest number of df., i.e. the MG2 factor model. The table shows parameter estimates for the loading parameters and the values of the NT and scaled chi-square goodness-of-fit tests. We can report that AR results for model MG2 show up in the similarity of regular (i.e. NT) and robust se in the case of loading parameters (not so similar in the case of variance parameters) and regular and robust chi-square goodness-of-fit test (including that obtained from AO analysis)[9]. Table 7 shows the results for model MG2-C bis (which passes the chi-square test at the 5% significance level), though no substantial differences can be reported in the estimates of parameters that are common to both MG2 and MG2-C bis models.

From the analysis of model MG2 in Table 6, we can report the generally high values of the loadings for variables V1 to V3, reflecting an adequate measurement model (the standardized values range from .5 to .99), whereas V2 (disposition to pay higher taxes) is the variable with the largest loadings. For V2, the standardized loading ranges across countries from .71 to .92. The loadings associated with V4 are smaller than the others, but we might interpret this loading as a regression coefficient, thus its low value does not deteriorate the measurement model for

[8] The estimates of variance are available from the authors upon request. For reasons of space, we chose not to include these data here.

[9] The robust and non-robust se differ by less than .001, and the correlation across countries among the NT chi-square, the scaled and AO se are above .997, with the mean (across countries) values of these statistics differing by less than .1.

F1. With respect to the goodness of fit of the model, the overall fit is highly significant, in accordance with highly significant chi-square values for the fit in each country. The inclusion of correlation effects in some countries decreases the chi-square goodness statistic dramatically. For example, in Germany-West, where the NT chi-square is 31.30 (d.f. = 2), the inclusion of a correlation between V3 (sacrifice standard of living) and V4 (participation in recycling) yields a chi-square value of 2.5[10]. This is not the case for the Netherlands, where the correlation needed is between V2 (pay higher taxes) and V4 (participation in recycling). For some countries, however, the model has a non-significant chi-value despite the high sample size. This is the case of Poland, Australia, Philippines, Russia, Ireland, Israel and Slovenia, where the NT chi-square is smaller than 2.5. Even though all the regression coefficients are significant at the 5 % level, for some countries the effect is small enough for it not to be important. The countries with a low effect between Attitude and Behavior are Germany-West, Japan, Russia and Philippines, whose standardized regression coefficient is below .06; on the other hand, countries like Northern Ireland, Bulgaria, Great Britain and Norway have standardized regression coefficients above .24. A multiple group model that passes the chi-square test is model MG2-C bis, whose estimation and tests results are shown in Table 7. No substantial difference can be seen when comparing the estimates of loading parameters resulting of Tables 6 and 7, despite the clear difference on the misfit of the model in terms of chi-square goodness-of-fit test of both MG2 and MG2-C bis models.

Despite violation of the factorial invariance hypothesis, multiple group models and MIMIC approach give comparable results. Table 8, showing the results for model MIMIC-C, offers an overview of the variation across countries (in relation to a baseline country, in this case Australia) of the mean levels of latent and manifest variables. We see, for example, that the Netherlands induces the largest positive country effect on F1 (.405), while the Czech Republic has the largest negative effect on F1 (-.501). With regard to the direct effect of country on variables, we see a large negative value of Russia on V4 (-1.36) while there is a large positive value (.359 and .317) in both Germany-West and East. Bulgaria, Israel and Northern Ireland also induce a large negative effect on V4[11].

Conclusions

The paper has considered alternative approaches to the analysis of group differences in both the mean levels of manifest and latent variables, and in their mutual relationship. Continuous versus ordinal methods, single group versus multiple group methods, NT versus asymptotic robust methods have been applied to ISSP-1993 survey data. The specific features of the data file are the large number of groups, the large sample size in each group, and the marked non normality of observable variables.

For the ISSP-1993 Environmental data set, and the given SEM context, it was found that major substantive conclusions are fairly robust to the choice of alternative model strategies and

---

[10] In contrast with this marked change in the chi-square statistics, the (LISREL) GFI and AGFI are respectively .98 and .92 for the models without and with the correlation parameter.

[11] It would now be the task of more substantively oriented researchers to interpret these findings in the ISSP-1993 Environment data.

estimation methods. Indeed alternative approaches to SEM analysis coincide in rejecting: 1) a null hypothesis of homogeneity across countries of variances of unique and common factors on “attitude” to and “behavior” concerning environmental issues; 2) the hypothesis of invariance across country of measurement loadings and the regression effect of “attitude” to “behavior”. Furthermore, for some countries, a direct correlation between variable V3 (disposition to cutting standard of living) and V4 (actual recycling behavior), not accounted for by the common factor F, was detected empirically. These findings are clearly visible from the chi-square values shown in Table 3.

The differences between countries in the measurement structure of variables V1-V3 when measuring F1 (positive disposition toward protecting environment), and the variation across countries of the regression effect of F1 to V4, are displayed in Table 7 (which gives results for a multiple group model MG2-C bis). The direct effects of the countries on both the level of the common factor F1 and the manifest variables V1-V4 were isolated and quantified: these effects are shown in Table 8 (which shows results for the single group model MIMIC-C).

The MIMIC model was found useful for summarizing the effects of countries on both latent and manifest variables, despite the violation of the assumption of homogeneity across groups of variances of unique and common factors.

Minimal differences were found between NT and robust se in the case of CO analysis when estimating loadings and country effects for those models in which the MC of AR applies. Although we were involved with a very large data set, the chi-square goodness-of-fit test was useful for model assessment.

References

Ajzen, I., & Fishbein, M. (1982). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology, 30*, 557-576.

Coenders, G., Saris, W.E., & Satorra, A. (1997). Alternative approach to structural modeling of ordinal data: a Monte Carlo study, *Structural Equation Modeling, 4*, 261-282.

Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika, 40*, 5-32.

Finch, J. F., West, S.G., & MacKinnon D.P. (1997). Effects of sample size and nonnormality on the estimation of mediated effects in latent variable models, *Structural Equation Modeling, 4*, 87-107.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: an introduction to theory and research*. Reading, Mass: Addison-Wesley.

Green, S.B., Akey T.M., Fleming K.K. Hershberger S.L., & Marquis, J.G. (1997). Effect of the number of scale points on chi-square fit indices in confirmatory factor analysis, *Structural Equation Modeling, 4*, 108-120.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36,* 409-426.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of American Statistical Association, 10*, 631-639.

Jöreskog, K. G., & Sörbom D. (1996). *LISREL 8: User's reference guide*. Chicago, Scientific Software International.

Kaplan, D. (1991). The behavior of three weighted least squares estimators for structured means analysis with non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 44*, 333-346.

Kaplan, D., & George R. (1995). A study of the power associated with testing factor mean differences under violation of factorial invariance, *Structural Equation Modeling, 2,* 101-118.

Meredith, W. (1993) Measurement invariance, factor analysis and factorial invariance, *Psychometrika*, *58*, 525-543.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49*, 115-132.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, *54*, 557-585.

Muthén, L.K., & Muthén B.O. (1998). *Mplus User's Guide*. Los Angeles.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation. *Psychometrika*, *44*, 443-460.

Pieters, R. G. M. (1988). Attitude-behavior relationships. In W. F. Van Raaij, G. M. Van Veldhoven & K. E. Wärneryd (Eds.), *Handbook of Economic Psychology* (pp. 144-204). Dordrecht, The Netherlands, Kluwer Academic Publishers.

Saris W.E., den Ronden, J., & Satorra A. (1987). Structural modeling by example: applications in educational. In P.F. Cuttance & J.R. Escob (Eds.), *Sociological, and Behavioural Research* (pp. 202-220). Cambridge University Press

Satorra A. (1992). Asymptotic robust inferences in the analysis of mean and covariance structures. In P. V. Marsden (Ed.), Sociological Methodology 1992 (pp. 249-278). Oxford & Cambridge, MA: Basil Blackwell.

Satorra A. (1993). Asymptotic robust inferences in multi-sample analysis of augmented moment structures. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate Analysis: Future Directions II* (pp. 211-229). Elsevier: Amsterdam.

Satorra, A. and P.M. Bentler (1994). Corrections to Test Statistics and standard errors in covariance structure analysis. In: Alexander von Eye and C.C. Clogg (Eds.). *Latent variables analysis: applications to developmental research* (pp. 399-419). SAGE Publications, Inc: Thousand Oaks, CA.

Satorra, A. & Bentler, P.M. (1999). A scaled difference chi-square test statistic for moment structure analysis. *UPF Economics Working Paper 412*, Barcelona

Sörbom, D. (1974). A general method for studying difference in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology, 27,* 229-239.

Appendix 1. Variables used in the study as they appear in the ISSP-1993 questionnaire

Questions	Responses
Protect enviro: pay much higher prices Q.8a How willing would you be to pay much higher prices in order to protect the environment?	1. Very willing[*] 2. Fairly willing 3. Neither willing nor unwilling 4. Fairly unwilling 5. Very unwilling 8. Can't choose, don't know 9. NA, refused
Protect enviro: pay much higher taxes Q.8b And how willing would you be to pay much higher taxes in order to protect the environment?	
Protect enviro: cut your standard of living Q.8c And how willing would you be to accept cuts in your standard of living in order to protect the environment?	

Effort: sort glass for recycling Q.19a How often do you make a special effort to sort glass or tins or plastic or newspapers and so on for recycling?	1. Always 2. Often 3. Sometimes 4. Never 5. Recycling not available where I live 8. Can't choose, don't know 9. NA, refused

[*] Variables Q. 8a – Q. 8c and Q. 19a correspond in our analysis to V1 - V4 respectively, but their response scales are reversed (i.e. 1 to 5 of Q. 8a – Q. 8c are 5 to 1 of V1 - V3; 1 to 4 of Q. 19a are 4 to 1 of V4).

TABLE 1. Model Equations

Measurement Equations[12]:

	INT	F1: Attitude	F2: Behavior	Ei
V1 Pay higher Price	0	1	0	*
V2 Pay higher Taxes	0	*	0	*
V3 Cut living standards	0	*	0	*
V4 Behavior	0	0	1	0

Structural Equations:

	INT	F1: Attitude	Di
F1: Attitude	0	*	0
F2: Behavior	0	*	*

[12] Asterisks denote free parameters to be estimated

TABLE 2. Models considered in the study

Across group constraints	Group Design: Multiple Group: Cov.	Group Design: Multiple Group: Mean-Cov.	Group Design: Single Group: MIMIC	Group Design: Single Group: One Factor Model
Loadings & Variances	MG0	MMG0	MIMIC-A MIMIC-B MIMIC-C	SG-A _ SG-C
Loadings	MG1	MMG1	_	_
Selected set of Loadings	MG12	MMG12 MMG12-C bis MMG12-C	_	_
No constraints	MG2 MG2-C bis MG2-C	_	_	_

TABLE 3. Overview of chi-square statistics for different models

Across group constraints	Group Design: Multiple Group – Cov.	Multiple Group – Mean-Cov.	Single Group – MIMIC	Single Group – One Factor Model
Loadings & Variances	**MG0** (d.f.=212)	**MMG0** (d.f. = 275)	**MIMIC-A** (d.f. = 65)	**SG-A** (d.f. = 2)
	3486.35 [a]	10466.35 [a]	7059.30 [a]	242.21 [a]
	3025.62 [b]	-	6887.53 [b]	228.84 [b]
	3439.25 [c]	-	-	223.31 [c]
	657.21 [d]	-	2318.08 [d]	214.84 [d]
	494.05 [e]	-	1307.64 [e]	215.39 [e]
	574.68 [f]	-	-	214.86 [f]
	_	_	**MIMIC-B** (d.f. = 12)	_
			74.16 [a]	
			77.71 [b]	
			-	
			33.64 [d]	
			29.25 [e]	
	_	_	**MIMIC-C** (d.f. = 11)	**SG-C** (d.f. = 1)
			34.72 [a]	34.57 [a]
			36.77 [b]	33.15 [b]
			-	33.19 [c]
			25.04 [d]	32.38 [d]
			25.28 [e]	32.38 [e]
			-	32.38 [f]
Loadings	**MG1** (d.f. = 107)	**MMG1** (d.f. = 170)	_	_
	473.65 [a]	7234.53 [a]		
	442.64 [b]	-		
	442.59 [c]	-		
	273.77 [d]	-		
	303.43 [e]	-		
	316.69 [f]	-		
Selected set of Loadings	**MG12** (d.f. = 87)	**MMG12** (d.f. = 143)	_	_
	220.74 [a]	407.36 [a]		
	205.24 [b]	-		
	207.92 [c]	-		
	196.65 [d]	-		
	216.51 [e]	-		
	217.71 [f]	-		
		MMG12-C bis (d.f. = 135)		
		329.01 [a]		
		MMG12-C (d.f. = 121)		
		311.90 [a]		
No constraints	**MG2** (d.f. = 44)	_	_	_
	149.61 [a]			
	141.34 [b]			
	142.56 [c]			
	125.82 [d]			
	146.69 [e]			
	139.58 [f]			
	MG2-C bis (d.f. = 36)			
	63.45 [a]			
	59.46 [b]			
	60.95 [c]			
	52.09 [d]			
	54.29 [e]			
	58.14 [f]			
	MG2-C (d.f. = 22)			
	50.67 [a]			
	48.30 [b]			
	48.09 [c]			
	39.04 [d]			
	42.86 [e]			
	43.58 [f]			

[a] CO – NT chi-square statistics
[b] CO – Scaled chi-square statistics
[c] CO – ADF chi-square statistics
[d] OR – Scaled chi-square statistics
[e] OR – Robust chi-square statistics
[f] OR – WLS chi-square statistics

TABLE 4. NT estimates of loading and variance parameters for different models

	Parameter Estimates									
	Loadings				Variances and Covariance					
Models	V1,F1	V2,F1	V3,F1	V4,F1	F1,F1	E1,E1	E2,E2	E3,E3	E4,E4	E3,E4
MG0	1	1.147	0.938	0.201	0.761	0.482	0.406	0.666	0.908	_
		0.011[a]	0.010	0.008	0.013	0.007	0.009	0.008	0.009	
		0.012[b]	0.011	0.008	0.013	0.010	0.011	0.011	0.007	
MG1	1	1.132	0.945	0.196	_	_	_	_	_	_
		0.010	0.009	0.008						
		0.011	0.010	0.008						
MMG0	1	1.141	0.943	0.224	0.754	0.503	0.428	0.708	1.136	_
		0.011	0.010	0.009	0.012	0.008	0.009	0.009	0.011	
MMG1	1	1.123	0.949	0.236	_	_	_	_	_	_
		0.010	0.009	0.009						
MIMIC-A	1	1.139	0.944	0.225	0.753	0.503	0.430	0.706	1.135	_
		0.011	0.010	0.009	0.012	0.008	0.009	0.009	0.011	
		0.012	0.010	0.009	0.013	0.010	0.012	0.011	0.011	
MIMIC-B	1	1.146	0.936	0.201	0.761	0.481	0.405	0.666	0.907	_
		0.011	0.010	0.008	0.013	0.007	0.009	0.008	0.009	
		0.012	0.011	0.008	0.013	0.010	0.011	0.011	0.007	
MIMIC-C	1	1.150	0.934	0.190	0.760	0.482	0.400	0.670	0.911	0.038
		0.011	0.010	0.008	0.013	0.008	0.009	0.008	0.009	0.006
		0.012	0.010	0.009	0.013	0.010	0.011	0.011	0.007	0.006
SG-A	1	1.146	0.941	0.218	0.806	0.505	0.419	0.712	1.137	_
		0.011	0.010	0.009	0.013	0.008	0.009	0.009	0.011	
		0.012	0.010	0.009	0.014	0.011	0.012	0.011	0.007	
SG-C	1	1.156	0.936	0.192	0.803	0.509	0.406	0.722	1.146	0.100
		0.011	0.010	0.009	0.013	0.008	0.009	0.009	0.011	0.007
		0.012	0.011	0.009	0.014	0.011	0.012	0.011	0.007	0.007

[a] NT se
[b] Robust se

TABLE 5. CO-OR NT standardized loading parameters for different models

	Parameter Estimates									
	Loadings				Variances and Covariance					
Models	**V1,F1**	**V2,F1**	**V3,F1**	**V4,F1**	**F1,F1**	**E1,E1**	**E2,E2**	**E3,E3**	**E4,E4**	**E3,E4**
MG0	0.782 [a]	0.843	0.708	0.181	1.000	0.388	0.288	0.498	0.966	_
	0.835 [b]	0.885	0.752	0.200	1.000	0.303	0.217	0.434	0.960	_
MG1	.	.	.	.	.	.	.	.	.	_
MMG0	0.774	0.834	0.697	0.179	0.968	0.401	0.304	0.514	0.968	_
MMG1	.	.	.	.	.	.	.	.	.	_
MIMIC-A	0.785	0.843	0.710	0.186	0.932	0.384	0.289	0.496	0.966	_
	0.827	0.887	0.748	0.190	0.934	0.316	0.214	0.440	0.964	–
MIMIC-B	0.798	0.860	0.716	0.170	0.913	0.367	0.274	0.467	0.773	_
	0.844	0.901	0.756	0.182	0.914	0.294	0.203	0.407	0.736	–
MIMIC-C	0.797	0.863	0.714	0.160	0.913	0.368	0.270	0.460	0.775	0.029
	0.843	0.903	0.754	0.174	0.914	0.296	0.199	0.409	0.739	0.031
SG-A	0.784	0.847	0.707	0.181	1.000	0.385	0.283	0.499	0.967	_
	0.828	0.887	0.747	0.187	1.000	0.315	0.212	0.442	0.965	–
SG-C	0.782	0.852	0.703	0.159	1.000	0.388	0.274	0.506	0.975	0.077
	0.825	0.893	0.743	0.166	1.000	0.319	0.203	0.448	0.972	0.087

[a] CO - NT
[b] OR - NT

TABLE 6. MG2: CO-NT parameter estimates and chi-square statistics

	Sample	Chi-square (d.f.=2) [a]			Parameter Estimates							
					Loadings				Standardized Loadings			
Countries	**Size**	**ADF**	**NT**	**Scaled**	**V1,F1**	**V2,F1**	**V3,F1**	**V4,F1**	**V1,F1**	**V2,F1**	**V3,F1**	**V4,F1**
Australia	1613	0.01	0.01	0.01	1	1.13 *	0.92	0.24	0.82	0.85	0.73	0.21
Germany-West	927	29.48	31.30	33.16	1	0.95	0.74	0.15	0.86	0.81	0.68	0.20
Germany-East	949	12.83	12.47	14.09	1	0.91	0.81	0.06	0.84	0.82	0.69	0.07
Great Britain	1091	3.98	3.80	3.90	1	1.19	0.97	0.30	0.82	0.91	0.73	0.25
Northern Ireland	600	7.05	6.61	5.82	1	1.11	1.03	0.37	0.79	0.82	0.75	0.32
United States	1364	4.41	4.72	4.50	1	1.16	0.97	0.24	0.81	0.87	0.74	0.20
Hungary	882	5.15	5.41	5.11	1	1.17	0.82	0.23	0.72	0.92	0.68	0.17
Italy	761	10.40	10.99	10.51	1	1.30	0.97	0.26	0.72	0.78	0.66	0.18
Ireland	681	2.28	2.35	2.25	1	0.99	1.08	0.25	0.71	0.71	0.75	0.21
Netherlands	1701	11.31	13.00	11.00	1	1.44	1.14	0.22	0.75	0.86	0.71	0.20
Norway	1110	3.60	3.29	3.21	1	1.01	0.90	0.26	0.84	0.83	0.72	0.24
Czech Republic	495	4.51	4.67	4.75	1	0.95	0.86	0.13	0.77	0.75	0.66	0.11
Slovenia	587	2.25	2.47	2.34	1	1.05	0.89	0.25	0.89	0.88	0.79	0.24
Poland	754	0.00	0.00	0.00	1	1.06	0.92	0.25	0.81	0.86	0.79	0.24
Bulgaria	564	4.75	5.91	5.11	1	1.05	0.96	0.20	0.90	0.92	0.85	0.30
Russia	760	2.09	2.32	1.81	1	1.22	0.80	0.07	0.78	0.92	0.60	0.08
New Zealand	1137	13.19	14.78	14.08	1	1.21	1.09	0.21	0.77	0.83	0.76	0.17
Canada	1274	9.17	9.14	9.46	1	1.20	0.98	0.21	0.77	0.81	0.71	0.18
Philippines	1165	0.74	0.75	0.75	1	1.23	0.90	0.09	0.67	0.83	0.61	0.08
Israel	761	1.92	2.14	1.94	1	1.35	1.05	0.29	0.73	0.87	0.70	0.23
Japan	1198	6.80	6.76	7.06	1	1.27	0.81	0.06	0.66	0.81	0.53	0.05
Spain	888	6.47	6.59	6.37	1	1.08	1.01	0.12	0.88	0.91	0.87	0.09

*All estimates are significant at 5% level

[a] These are chi-square for single sample analysis. The overall chi-square for this model is reported in Table 3

TABLE 7. MG2-C bis: CO-NT parameter estimates and chi-square statistics

				Parameter Estimates									
	Chi-square [a]			Loadings and Covariance					Standardized Loadings and Covariance				
Countries	**ADF**	**NT**	**Scaled**	**V1,F1**	**V2,F1**	**V3,F1**	**V4,F1**	**E3,E4**	**V1,F1**	**V2,F1**	**V3,F1**	**V4,F1**	**E3,E4**
Australia	0.01	0.01	0.01	1	1.13 *	0.92	0.24	0	0.82	0.85	0.73	0.21	0
Germany-West	_ [b]	2.50	2.87	1	0.96	0.74	0.13	0.12	0.87	0.82	0.67	0.17	0.19
Germany-East	_	0.22	0.26	1	0.92	0.81	0.04	0.09	0.84	0.82	0.69	0.05	0.12
Great Britain	_	0.11	0.10	1	1.19	0.97	0.29	0.05	0.82	0.91	0.73	0.25	0.06
Northern Ireland	7.05	6.61	5.82	1	1.11	1.03	0.37	0	0.79	0.82	0.75	0.32	0
United States	4.41	4.72	4.50	1	1.16	0.97	0.24	0	0.81	0.87	0.74	0.20	0
Hungary	5.15	5.41	5.11	1	1.17	0.82	0.23	0	0.72	0.92	0.68	0.17	0
Italy	3.64	3.64	3.48	1	1.32	0.97	0.21	0.12	0.72	0.78	0.66	0.18	0.11
Ireland	2.28	2.35	2.25	1	0.99	1.08	0.25	0	0.71	0.71	0.75	0.21	0
Netherlands	_	6.41	5.64	1	1.46	1.14	0.21	0.04	0.75	0.86	0.71	0.20	0.07
Norway	3.60	3.29	3.21	1	1.01	0.90	0.26	0	0.84	0.83	0.72	0.24	0
Czech Republic	4.51	4.67	4.75	1	0.95	0.86	0.13	0	0.77	0.75	0.66	0.11	0
Slovenia	2.25	2.47	2.34	1	1.05	0.89	0.25	0	0.89	0.88	0.79	0.24	0
Poland	0.00	0.00	0.00	1	1.06	0.92	0.25	0	0.81	0.86	0.79	0.24	0
Bulgaria	4.75	5.91	5.11	1	1.05	0.96	0.20	0	0.90	0.92	0.85	0.30	0
Russia	2.09	2.32	1.81	1	1.22	0.80	0.07	0	0.78	0.92	0.60	0.08	0
New Zealand	0.01	0.01	0.01	1	1.22	1.09	0..18	0.09	0.77	0.83	0.76	0.17	0.13
Canada	3.35	3.15	3.37	1	1.21	0.98	0.19	0.06	0.77	0.81	0.71	0.18	0.08
Philippines	0.74	0.75	0.75	1	1.23	0.90	0.09	0	0.67	0.83	0.61	0.08	0
Israel	1.92	2.14	1.94	1	1.35	1.05	0.29	0	0.73	0.87	0.70	0.23	0
Japan	_	0.03	0.04	1	1.28	0.81	0.05	0.07	0.66	0.81	0.53	0.05	0.08
Spain	6.47	6.59	6.37	1	1.08	1.01	0.12	0	0.88	0.91	0.87	0.09	0

*All estimates are significant at 5% level

[a] These are chi-square for single sample analysis. The d.f. are 1 or 2 depending respectively on whether the correlation is present or not. The overall chi-square for this model is reported in Table 3

[b] ADF did not converge

TABLE 8. MIMIC-C: NT Parameter estimates

Loadings, Variances and Covariance

	F1	V1	V2	V3	V4
F1	0.760*	1	1.150*	0.934*	0.190*
E		0.482*	0.400*	0.670*	0.911*
E3,E4				0.038 (0.0048)[a]	

Country Effects

Countries	F1	V1	V2	V3	V4
Australia [b]	-	-	-	-	-
Germany-West	-0,068	0	0	0.387*	0.359*
Germany-East	-0.580*	0	0.152*	0.513*	0.317*
Great Britain	-0.018	0	0.075*	-0,288*	-0.595*
Northern Ireland	-0.123*	0	0	-0.338*	-1.021*
United States	0.031	0	0	-0.208*	-0.181*
Hungary	-0.652	0.206*	0	-0.393*	-0.710*
Italy	0.152	0.152*	-0.346*	0	-0.488*
Ireland	-0.627	0.416*	0	0	-0.713*
Netherlands	0.405*	0	-0.256*	-0.207*	0.124*
Norway	0.010	0	-0.101*	0.264*	-0.667*
Czech Republic	-0.501*	0	0	-0.178*	-0.390*
Slovenia	0.021	0	0.121*	0	-0.418*
Poland	-0.169*	0	0.106*	-0.210*	-0.766*
Bulgaria	0.014	0	0.171*	0	-1.360*
Russia	-0.059	0.117*	0.372*	0	-1.418*
New Zealand	-0.028	0.081*	0	-0.071*	-0.256*
Canada	0.102*	0	-0.219*	0	-0.070*
Philippines	-0.372*	-0.246*	0	0.091*	-0.446*
Israel	0.125*	0	-0.219*	-0.592*	-1.242*
Japan	-0.049	0	0.099*	0	0.099*
Spain	0.167*	-0.142*	0	0	-0.821*

[a] Estimated covariance between residual errors. The corresponding correlation is in parenthesis

[b] Australia acts as baseline country

* Significant at 5% level