THURSTONIAN COVARIANCE AND CORRELATION STRUCTURES FOR MULTIPLE JUDGMENT PAIRED COMPARISON DATA

IE Working Paper **WP 04 / 03** **30 / 04 / 2003**

Alberto Maydeu Olivares

Instituto de Empresa
Marketing Dept.
C / María de Molina 11-15,
28006 – Madrid
España
Alberto.Maydeu@ie.edu

Abstract
Thurstone's (1927) is not a proper model for multiple judgment paired comparison data as it assigns zero probabilities to all intransitive patterns. To obtain a proper model, Takane (1987) extended Thurstone's model by adding a vector of pair specific random errors. We investigate an unrestricted Thurstone-Takane model when (a) the variances of the paired specific errors are assumed to be equal, and (b) when they are assumed to be unequal.
We also consider a new model, an unrestricted Thurstonian mean and correlation structure model. As the Thurstone-Takane model is a mean and covariance structure model, and its covariance structure is not scale invariant, Thurstonian correlation and covariance structure models are not equivalent, except in some restricted cases. In particular, this correlation structure model is equivalent to the mean and covariance structure model considered in Maydeu-Olivares (2001). Yet, their substantive interpretation differ.

Keywords
MPLUS, goodness of fit, scale invariance, limited information estimation, structural equation models, personnel assessment, comparative data, random utility models, Case III model, Case V model

1. Introduction

Given a set of stimuli and a sample of respondents, a multiple judgment paired comparison experiment consists in presenting all possible pairs of stimuli to each respondent asking them to choose one object within each pair. These experiments have a long tradition and they are particularly popular in areas such as psychophysics and consumer behavior. Several models have been proposed for these type of data (see David, 1988), one of the oldest being proposed by Thurstone. Thurstone's (1927) model is characterized by three assumptions: (a) whenever a pair of stimuli is presented to a respondent it elicits a continuous preference (utility function, or in Thurstone's terminology, *discriminal process*) for each stimulus; (b) within a pair, the stimulus whose continuous preference is largest will be preferred by the respondent; (c) the continuous preferences are normally distributed in the population.

In a multiple judgment experiment where the responses to n stimuli are to be modeled, one can observe $2^{\tilde{n}}$ binary patterns, where $\tilde{n} = \binom{n}{2} = \frac{n(n-1)}{2}$. Of these patterns, $n!$ are transitive, meaning that given the binary patterns one can rank order the stimuli, and the rest intransitive. Maydeu-Olivares (1999) showed that the model proposed by Thurstone is not a proper model for multiple judgment paired comparison data as it assigns zero probabilities to all intransitive patterns. Takane must have been aware of this, as in 1987 he proposed adding a vector of pair specific random errors to Thurstone's (1927) model. This extension of Thurstone's model, henceforth referred in this paper as the Thurstone-Takane model, is a proper model for multiple judgment paired comparisons as it assigns non-zero probabilities to all paired comparisons patterns. Takane's (1987) crucial contribution, however, was largely programmatic, and he provided neither identification restrictions nor empirical examples. Maydeu-Olivares (2001) proposed a set of identification restrictions for the Thurstone-Takane model and investigated the performance of limited information methods (Muthén, 1978, 1993; Muthén, du Toit & Spisic, in press) to estimate the model. He found that limited information methods have an excellent small sample behavior in estimating and testing this model even with large number of stimuli. Yet, he also reported that in applications often improper solutions were obtained in which estimates for the variances of the pair specific errors became negative.

The purpose of the present contribution is threefold.

1) We show that these improper solutions occur because the identification restrictions introduced by Maydeu-Olivares (2001) in the Thurstone-Takane

model are unnecessarily restrictive. Consequently, we provide an alternative set of minimal identification restrictions to identify the Thurstone-Takane model.

2) We investigate by means of a simulation study the small sample performance of limited information methods to recover a properly specified Thurstone-Takane model.
3) We propose a new model for multiple judgment paired comparisons data based on Thurstone's (1927) original proposals. That is, we introduce a solution –different from Takane's (1987)- to the problem of specifying a Thurstonian model for multiple judgment pair comparison data that assigns non-zero probabilities to all paired comparisons patterns.

To do so, we shall first review the restrictions imposed by Thurstone's model and the Thurstone-Takane model on the means, variances and covariances of the pairwise differences among the continuous preferences assumed by Thurstone (1927). Next, we discuss the identification of the Thurstone-Takane model, showing that Maydeu-Olivares' (2001) identification restriction is unnecessarily restrictive. Next, we perform a simulation study to investigate how well limited information estimation methods recover an unrestricted Thurstone-Takane model under two conditions: equal and unequal variances for the paired specific errors. Then, we introduce a new Thurstonian model for multiple judgment paired comparisons data. Unlike the Thurstone-Takane model, our model introduces restrictions only on the means and correlations -the variances are left unspecified- of the pairwise differences among the continuous preferences assumed by Thurstone (1927). We conclude our presentation with numerical examples to illustrate the behavior of Thurstonian covariance and correlation structure models in fitting several actual datasets.

2. Thurstonian Covariance Structure Models for Multiple Judgment Paired Comparison Data

2.1 Thurstone's model

Consider a set of n stimuli and a random sample of N individuals from the population we wish to investigate. In a multiple judgment paired comparison experiment $\tilde{n} = \binom{n}{2} = \frac{n(n-1)}{2}$ pairs of stimuli are constructed and each pair is presented to each individual in the sample. We shall denote by y_l the outcomes of each paired comparison

$$y_{lj} = \begin{cases} 1 & \text{if} \quad \text{subject } j \text{ chooses object } i \\ 0 & \text{if} \quad \text{subject } j \text{ chooses object } i' \end{cases} \qquad l = 1,\, ...,\, \tilde{n};\;\; j = 1,\, ...,\, N \tag{1}$$

where $l \equiv (i,i')$,(i = 1,...,n - 1; i' = i+1,...,n). Let t_{ij} denote subject j's unobserved continuous preference for object i. According to Thurstone's (1927) model: (a) the preferences **t** are normally distributed in the population, and (b) a subject will choose object i if $t_{ij} \geq t_{i'j}$, otherwise s/he will choose object i'.

Thurstone (1927) proposed performing the following linear transformation on the set of unobserved preferences

$$y_l^* = t_i - t_{i'} \tag{2}$$

Then, (b) may be alternatively expressed as

$$y_{lj} = \begin{cases} 1 & \text{if} \quad y_{lj}^* \geq 0 \\ 0 & \text{if} \quad y_{lj}^* < 0 \end{cases} \tag{3}$$

In matrix notation, Thurstone's model can be expressed as follows: Let $\mathbf{t} \sim N\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right)$, then $\mathbf{y}^* = \mathbf{A}\,\mathbf{t}$. **A** is a $\tilde{n} \times n$ design matrix where each column corresponds to one of the stimuli, and each row to one of the paired comparisons. For example, when n = 4, **A** is

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \tag{4}$$

It then follows that the probability of any paired comparisons pattern under Thurstone's (1927) model is

$$\Pr\left(\bigcap_{l=1}^{\tilde{n}} y_l\right) = \int \underset{\mathbf{R}}{\cdots} \int \phi_{\tilde{n}}\left(\mathbf{y}^* : \boldsymbol{\mu}_{y^*}, \boldsymbol{\Sigma}_{y^*}\right) d\mathbf{y}^* \tag{5}$$

$$\boldsymbol{\mu}_{y^*} = \mathbf{A}\boldsymbol{\mu}_t \qquad\qquad \boldsymbol{\Sigma}_{y^*} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' \tag{6}$$

where $\phi_{\tilde{n}}(\bullet)$ denotes a $\tilde{n}$-variate normal density, and $\mathbf{R}$ is the multidimensional rectangular region formed by the product of intervals

$$R_l = \begin{cases} (0,\infty) & \text{if} \quad y_l = 1 \\ (-\infty,0) & \text{if} \quad y_l = 0 \end{cases}. \tag{7}$$

Now, since $\mathbf{A}$ is of rank n - 1, $\boldsymbol{\Sigma}_{y^*}$ has rank $n-1$ and Thurstone's model assigns zero probabilities to all intransitive patterns (Maydeu-Olivares, 1999). Thus, Thurstone's model is not a plausible model for multiple judgment paired comparisons data, but it may be a suitable model for ranking data.

2.2 Thurstone-Takane's model

Takane's (1987) key contribution was to add a random error e_l to each paired comparison (2). He assumed that these errors were independent of each other and independent of the unobserved continuous preferences $\mathbf{t}$ so that

$$\begin{pmatrix} \mathbf{t} \\ \mathbf{e} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\mu}_t \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_t & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}^2 \end{pmatrix} \right) \tag{8}$$

where $\boldsymbol{\Omega}^2$ is a diagonal matrix with elements ω_l^2. Then, $\mathbf{y}^* = \mathbf{A}\,\mathbf{t} + \mathbf{e}$, and from (3), we find that the probability of any paired comparisons pattern under the Thurstone-Takane model is given also by (5) and (7) but now

$$\boldsymbol{\mu}_{y^*} = \mathbf{A}\boldsymbol{\mu}_t \qquad \boldsymbol{\Sigma}_{y^*} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \boldsymbol{\Omega}^2 \tag{9}$$

Now, since the observed data is binary, the pattern probabilities (5) are unchanged when we standardize $\mathbf{y}^*$ using

$$\mathbf{z}^* = \mathbf{D}\left(\mathbf{y}^* - \boldsymbol{\mu}_{y^*}\right) \qquad \mathbf{D} = \left(\text{Diag}\left(\boldsymbol{\Sigma}_{y^*}\right)\right)^{-\frac{1}{2}}. \tag{10}$$

Furthermore, letting μ_l^* be an element of $\boldsymbol{\mu}_{y^*}$ and $\sigma_l^{2^*}$ be a diagonal element of $\boldsymbol{\Sigma}_{y^*}$ we have that at $y_l^* = 0$, $\tau_l := \frac{-\mu_l^*}{\sqrt{\sigma_l^{2^*}}}$. As a result, $\boldsymbol{\mu}_{z^*} = \mathbf{0}$ and

$$\boldsymbol{\tau} = -\mathbf{D}\boldsymbol{\mu}_{y^*} = -\mathbf{D}\mathbf{A}\boldsymbol{\mu}_t \qquad \boldsymbol{P}_{z^*} = \mathbf{D}\boldsymbol{\Sigma}_{y^*}\mathbf{D} = \mathbf{D}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \boldsymbol{\Omega}^2\right)\mathbf{D} \tag{11}$$

We use $\boldsymbol{P}$ in (11) to indicate that the covariance matrix of $\mathbf{z}^*$ has ones along its diagonal. Furthermore, $\boldsymbol{\tau}$ denotes a vector of thresholds, and the off-diagonal elements of $\boldsymbol{P}_{z^*}$ are tetrachoric correlations.

Therefore, the Thurstone-Takane model defined by equations (5), (7), and (9) can be equivalently rewritten as

$$\Pr\left(\bigcap_{l=1}^{\tilde{n}} y_l\right) = \int \underset{\breve{\mathbf{R}}}{\cdots} \int \phi_{\tilde{n}}\left(\mathbf{z}^* : \mathbf{0}, \boldsymbol{P}_{z^*}\right) d\mathbf{z}^* \tag{12}$$

where $\breve{\mathbf{R}}$ is the multidimensional rectangular region formed by the product of intervals

$$\breve{R}_l = \begin{cases} \left(\tau_l, \infty\right) & \text{if} \quad y_l = 1 \\ \left(-\infty, \tau_l\right) & \text{if} \quad y_l = 0 \end{cases} \tag{13}$$

and $\boldsymbol{\tau}$ and $\boldsymbol{P}_{z^*}$ are constrained as in (11).

Equations (11), (12), and (13) define in fact a class of models as $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ can be restricted in various ways. Takane (1987) provides an excellent overview of restricted Thurstonian models. We shall use the term *unrestricted* model when only minimal identification restrictions are imposed on $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$.

An interesting special case of the Thurstone-Takane class of models was proposed by Takane (1987) in which it is assumed that $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$.

2.3 Identification restrictions

From (11), (12), and (13) it is clear that to investigate the identification of any member of the Thurstone-Takane class it suffices to investigate the identification of the structure $\boldsymbol{\kappa}(\boldsymbol{\theta}) = \left(\boldsymbol{\tau}(\boldsymbol{\theta})', \boldsymbol{\rho}(\boldsymbol{\theta})'\right)'$, where $\boldsymbol{\theta}$ denotes the model parameters and $\boldsymbol{\rho}$ denotes the elements below the diagonal of $\boldsymbol{P}_{z^*}$ stacked onto a column vector.

We shall consider now the identification of an unrestricted model when $\boldsymbol{\Omega}^2$ is assumed to be diagonal. We first note that because of the comparative nature of the data it is necessary to set the location for the elements of $\boldsymbol{\mu}_t$ and the location for the elements in each of the rows (columns) of $\boldsymbol{\Sigma}_t$. Also, it is also necessary to set the location for the diagonal elements of $\boldsymbol{\Omega}^2$. Arbitrarily, we shall set $\mu_n = 0$, $\omega^2_{\tilde{n}} = 1$, and all the diagonal elements of $\boldsymbol{\Sigma}_t$

equal to one. Henceforth, we shall use $\boldsymbol{P}_t$ to indicate that $\boldsymbol{\Sigma}_t$ has ones along its diagonal.

Similar identification conditions can be established for other models within this class. For instance:

(a) When $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$, an unrestricted model is identified by letting $\mu_n = 0$, $\boldsymbol{\Sigma}_t = \boldsymbol{P}_t$, and $\omega^2 = 1$.

(b) A model where $\boldsymbol{\Sigma}_t$ is assumed to be a diagonal matrix (Thurstone's Case III model) is identified by letting $\mu_n = 0$ and $\omega_{\tilde{n}}^2 = 1$ if it is assumed that $\boldsymbol{\Omega}^2$ is diagonal, or by $\mu_n = 0$ and $\omega^2 = 1$ if it is assumed that $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$.

(c) Similarly, a model where it is assumed that $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$ (Thurstone's Case V model) is identified as in (b).

2.4 Maydeu-Olivares' (2001) identification restrictions

Maydeu-Olivares (2001) did not consider the standarization (10). Rather, he considered the problem of identifying the Thurstone-Takane class of models using equations (5), (7), and (9). Within this framework, since the variances in $\boldsymbol{\Sigma}_{y^*}$ may not be identified from the observed binary data, he simply suggested setting them equal to one by using

$$\boldsymbol{\Omega}^2 = \mathbf{I} - \mathrm{Diag}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'\right). \tag{14}$$

In addition, to identify an unrestricted model he suggested employing $\mu_n = 0$ and $\boldsymbol{\Sigma}_t = \boldsymbol{P}_t$ as we do here. With these identification restrictions, $\boldsymbol{\kappa}(\boldsymbol{\theta})$ is a linear function in the case of the unrestricted model. However, this solution is unnecessarily restrictive. It introduces $(\tilde{n} - 1)$ unnecessary restrictions. Also, Maydeu-Olivares (2001) reports it often yields improper solutions in applications with some of the diagonal elements of (14) becoming negative. Furthermore, if, upon encountering an improper solution, a proper solution is sought by imposing boundary constraints on the free parameters, a poorer fit will be obtained as $\boldsymbol{\kappa}(\boldsymbol{\theta})$ is linear.

3. Simulation study

We shall now investigate how well covariance structure models can be estimated in small samples. This is necessary as often only very small samples are available in paired comparisons applications, and the models under consideration are rather complex.

We shall use a limited information estimation approach due to Muthén (1978, 1993, Muthén, du Toit & Spisic, in press) to estimate these models. Let p_l and $p_{ll'}$ are the sample counterpart of $\pi_l = \Pr\left(y_l = 1\right)$ and $\pi_{ll'} = \Pr\left(y_l = 1, y_{l'} = 1\right)$, respectively, and let $\Phi_n(\bullet)$ denote a n-variate standard normal distribution function. Then, first each element of $\boldsymbol{\tau}$ is estimated separately using

$$\hat{\tau}_l = -\Phi_1^{-1}\left(p_l\right) \qquad l = 1, ..., \tilde{n} \tag{15}$$

Next, each element of $\boldsymbol{\rho}$ is estimated separately given the first stage estimates using

$$\hat{\rho}_{ll'} = \Phi_2^{-1}\left(p_{ll'} \left| -\hat{\tau}_l, -\hat{\tau}_{l'}\right.\right) \qquad l = 2, ..., \tilde{n}; \;\; l' = 1, ..., l - 1 \tag{16}$$

Finally, letting $\boldsymbol{\kappa} = \left(\boldsymbol{\tau}', \boldsymbol{\rho}'\right)'$, the model parameters $\boldsymbol{\theta}$ are estimated by minimizing

$$F = \left(\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}(\boldsymbol{\theta})\right)' \hat{\mathbf{W}} \left(\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}(\boldsymbol{\theta})\right) \tag{17}$$

where $\hat{\mathbf{W}} = \hat{\boldsymbol{\Xi}}^{-1}$ (WLS: Muthén, 1978), $\hat{\mathbf{W}} = \left(\text{Diag}\left(\hat{\boldsymbol{\Xi}}\right)\right)^{-1}$ (DWLS: Muthén, du Toit & Spisic, 1997), or $\hat{\mathbf{W}} = \mathbf{I}$ (ULS: Muthén, 1993), and $\boldsymbol{\Xi}$ denotes the asymptotic covariance matrix of $\sqrt{N}\left(\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}\right)$ which is computed as in Muthén (1978).

Standard errors for the parameter estimates are obtained using (Muthén, 1993)

$$\text{Acov}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{N}\mathbf{H}\boldsymbol{\Xi}\mathbf{H}' \qquad \mathbf{H} = \left(\boldsymbol{\Delta}'\mathbf{W}\boldsymbol{\Delta}\right)^{-1}\boldsymbol{\Delta}'\mathbf{W} \tag{18}$$

where $\boldsymbol{\Delta} = \dfrac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}'}$. Also, goodness of fit tests of the structural restrictions of the model $\boldsymbol{\kappa}(\boldsymbol{\theta})$ for the DWLS and ULS estimators can be obtained (Muthén, 1993) by scaling $T := N\hat{F}$ by its mean or adjusting it by its mean and variance so that it approximates a chi-square distribution as follows

$$T_s = \frac{r}{\text{Tr}[\mathbf{M}]}T \qquad T_a = \frac{\text{Tr}[\mathbf{M}]}{\text{Tr}\left[\mathbf{M}^2\right]}T \tag{19}$$

where $\mathbf{M} = \mathbf{W}(\mathbf{I} - \boldsymbol{\Delta}\mathbf{H})\boldsymbol{\Xi}$. T_s and T_a, denote the scaled (for mean) and adjusted (for mean and variance) test statistics. T_s is referred to a chi-square distribution with $r = \dfrac{\tilde{n}(\tilde{n}+1)}{2} - q$ degrees of freedom, where q is the number of mathematically independent parameters in $\boldsymbol{\theta}$.

T_a is referred to a chi-square distribution with $d = \frac{(\mathrm{Tr}[\mathbf{M}])^2}{\mathrm{Tr}[\mathbf{M}^2]}$ degrees of freedom.

Let $\mathbf{p}$ be the $\frac{\tilde{n}(\tilde{n}+1)}{2}$ vector of first and second order proportions employed in (15) and (16), and let $\boldsymbol{\pi}$ be its corresponding probabilities. Furthermore, let $\sqrt{N}\hat{\boldsymbol{\varepsilon}} := \sqrt{N}\left(\mathbf{p} - \boldsymbol{\pi}(\hat{\boldsymbol{\theta}})\right)$. Then, goodness of fit tests of the overall restrictions imposed by the model on the first and second order marginals of the contingency table $\boldsymbol{\pi}(\boldsymbol{\theta})$ can be obtained (Maydeu-Olivares, 2001) by scaling $\tilde{T} = N\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}$ using

$$\tilde{T}_s = \frac{r}{\mathrm{Tr}[\tilde{\mathbf{M}}]}\tilde{T} \qquad \tilde{T}_a = \frac{\mathrm{Tr}[\tilde{\mathbf{M}}]}{\mathrm{Tr}[\tilde{\mathbf{M}}^2]}\tilde{T}. \tag{20}$$

where $\tilde{\mathbf{M}} = \left(\mathbf{I} - \tilde{\boldsymbol{\Delta}}\boldsymbol{\Delta}\mathbf{H}\tilde{\boldsymbol{\Delta}}^{-1}\right)\boldsymbol{\Gamma}\left(\mathbf{I} - \tilde{\boldsymbol{\Delta}}\boldsymbol{\Delta}\mathbf{H}\tilde{\boldsymbol{\Delta}}^{-1}\right)'$, $\boldsymbol{\Gamma}$ denotes the asymptotic covariance matrix of $\sqrt{N}(\mathbf{p} - \boldsymbol{\pi})$, and $\tilde{\boldsymbol{\Delta}} = \frac{\partial \boldsymbol{\pi}}{\partial \boldsymbol{\kappa}'}$. $\tilde{T}_s$ and $\tilde{T}_a$ are to be referred to a chi-square distribution with r and $d = \frac{(\mathrm{Tr}[\tilde{\mathbf{M}}])^2}{\mathrm{Tr}[\tilde{\mathbf{M}}^2]}$ degrees of freedom, respectively.

Maydeu-Olivares (2001) investigated the small sample behavior of Muthén's procedures in estimating an unrestricted Thurstone-Takane's model with (14). He found that the ULS and DWLS estimators clearly outperformed the asymptotically efficient WLS estimator. The difference between ULS and DWLS was small, slightly favoring ULS. More specifically, Maydeu-Olivares (2001) found that using ULS to estimate an unrestricted Thurstonian model for 7 stimuli (21 binary variables being modeled), a sample size of 100 observations yielded parameters estimates with a relative bias $\frac{\overline{x}_{\hat{\theta}} - \theta_0}{\theta_0}$ at most of 2% (in absolute value), and standard errors with a relative bias $\frac{\overline{x}_{SE(\hat{\theta})} - sd_{\hat{\theta}}}{sd_{\hat{\theta}}}$ at most of 10% (also in absolute value). Furthermore, he found the mean and variance corrected statistics (T_a and $\tilde{T}_a$) to match rather closely its reference chi-square distribution, whereas the mean corrected statistics (T_s and $\tilde{T}_s$) showed an unacceptable behavior, rejecting the model all too often.

Nevertheless, the thresholds and tetrachoric correlations are a linear function of the parameters of the models considered by Maydeu-Olivares (2001). This may explain the excellent small sample behavior of the ULS procedure in his simulation. Larger sample sizes may be needed to estimate a correctly specified Thurstone-Takane model. To investigate this, we performed a simulation study under two conditions:

(a) Unrestricted Thurstone-Takane model with equal pair specific error variances, where $\boldsymbol{\tau} = -\mathbf{DA}\boldsymbol{\mu}_t$, and $\boldsymbol{P}_{z^*} = \mathbf{D}\left(\mathbf{A}\boldsymbol{P}_t\mathbf{A}' + \omega^2\mathbf{I}\right)\mathbf{D}$.

(b) Unrestricted Thurstone-Takane model with unequal pair specific error variances, where $\boldsymbol{\tau} = -\mathbf{DA}\boldsymbol{\mu}_t$, and $\boldsymbol{P}_{z^*} = \mathbf{D}\left(\mathbf{A}\boldsymbol{P}_t\mathbf{A}' + \boldsymbol{\Omega}^2\right)\mathbf{D}$.

To compare our results with those of Maydeu-Olivares (2001) we used the same estimation method (ULS in the third stage), the same number of stimuli (7), the same number of replications (1000), and the same true values used in Maydeu-Olivares (2001):

$$\boldsymbol{\mu}_t = \begin{pmatrix} 0.5 \\ 0 \\ -0.5 \\ 0 \\ 0.5 \\ -0.5 \\ 0 \end{pmatrix} \qquad \boldsymbol{P}_t = \begin{pmatrix} 1 & & & & & & \\ 0.8 & 1 & & & & & \\ 0.7 & 0.6 & 1 & & & & \\ 0.8 & 0.7 & 0.6 & 1 & & & \\ 0.8 & 0.7 & 0.6 & 0.8 & 1 & & \\ 0.7 & 0.6 & 0.8 & 0.7 & 0.6 & 1 & \\ 0.8 & 0.7 & 0.6 & 0.8 & 0.7 & 0.6 & 1 \end{pmatrix}. \tag{21}$$

In addition, in condition (b) we employed as true values $\boldsymbol{\Omega}^2 = \mathbf{I}$. To identify the models we let in both conditions $\mu_7 = 0$. Also, in (a) we let $\omega^2 = 1$ and in (b) we let $\omega^2_{21} = 1$.

The sample sizes investigated in condition (a) were 100 observations and 500 observations; and in condition (b) 500 and 1000 observations.

3.1 Simulation results when equal pair specific error variances are assumed

All replications reached convergence. For both sample sizes, all estimates for $\boldsymbol{\mu}_t$ showed a positive relative bias and all estimates for $\boldsymbol{P}_t$ showed a negative relative bias. There was no consistent trend in the sign of the relative bias of the estimated standard errors. In the smaller sample size, the largest parameter estimate and standard error relative biases were –2% and –7%, respectively. Thus, a sample size of 100 observations seems to be sufficient to obtain accurate parameter estimates and standard errors. This is truly remarkable given the degree of data sparseness and the complexity of the model. Table 1 summarizes the parameter estimates and standard errors results. Due to the large number of parameters being estimated, in this table we have pooled the results across parameter estimates having the same true value.

The results for the goodness of fit tests are shown in Table 2. We see in this table that none of the two statistics for assessing the structural restrictions of the model, T_s and T_a, closely match their reference chi-square distributions. Neither does the mean corrected statistic $\tilde{T}_s$ for assessing the overall restrictions of the model. However, the mean and variance corrected statistic $\tilde{T}_a$ for assessing the overall restrictions of the model shows a reasonable agreement to its reference distribution. A larger sample size is needed to obtain reliable tests of the structural restrictions for models of this size. Yet, a sample size of 100 observations suffices to obtain reliable tests of both the structural and overall restrictions in models for a smaller number of stimuli. We verified this in an additional simulation study with four stimuli.

Insert Tables 1 and 2 about here

When we used a sample size of 500 observations we obtained less variable parameter estimates, smaller relative biases, and now not only $\tilde{T}_a$ closely matched its reference distribution, but also T_s and T_a.

3.2 Simulation results when unequal pair specific error variances are assumed

Again, all replications converged. For both sample sizes, all estimates for $\boldsymbol{\mu}_t$ and $\boldsymbol{\Omega}^2$ showed a positive relative bias and all estimates for $\boldsymbol{P}_t$ showed a negative relative bias. All estimated standard errors showed a negative relative bias. When the sample size was 500, the largest parameter estimate relative biases for the elements of for $\boldsymbol{\mu}_t$ and $\boldsymbol{P}_t$ were 2% and –5%, respectively. However, relative biases for the elements of $\boldsymbol{\Omega}^2$ ranged between 6% and 28%. Also, the largest relative standard error biases for the elements of $\boldsymbol{\mu}_t$ was –5%. However, relative biases for the elements of $\boldsymbol{P}_t$ ranged between –10% and –17%, and for $\boldsymbol{\Omega}^2$ ranged between -7% and -22%., respectively. Clearly, this sample size is not large enough to obtain accurate estimates and standard errors for this particular model. However, as it can be seen in Table 2 it is large enough to reliably test the structural and overall restrictions imposed by this model, as the T_s and T_a and $\tilde{T}_a$ statistics closely match their reference distribution.

When we used a sample size of 1000 observations, the largest parameter estimate relative biases for $\boldsymbol{\mu}_t$, $\boldsymbol{P}_t$ and $\boldsymbol{\Omega}^2$ were 1%, -2% and 11%, respectively. Also, the largest standard error relative biases for $\boldsymbol{\mu}_t$, $\boldsymbol{P}_t$ and $\boldsymbol{\Omega}^2$ were -3%, -8% and -11%. This is remarkable, as we used the same starting seed in conditions (a) and (b). When the same data is estimated assuming that the variances are equal, 100 observations suffice to obtain accurate parameter estimates and standard errors, but 1000 observations are needed if the pair specific variances are to be estimated. Furthermore, inspecting the results in Table 1 we note that better results (more accurate and less variable parameter estimates are more accurate standard errors) are obtained with 100 observations when paired specific variances are assumed equal than with 500 observations when they are assumed to be unequal. Similarly, better results are obtained with 500 observations and equal variances than with 1000 observations and unequal variances.

When the pair comparison error variances are assumed unequal, the minimization (17) is performed with respect to $\boldsymbol{\vartheta} = \left(\boldsymbol{\mu}_t^{*\prime}, \boldsymbol{\rho}_t{}', \boldsymbol{\omega}'\right)'$, where $\boldsymbol{\mu}_t^* = \left(\mu_1, \cdots, \mu_6\right)'$, $\boldsymbol{\rho}_t = \left(\rho_{21}, \cdots, \rho_{76}\right)'$, and $\boldsymbol{\omega}^* = \left(\omega_1, \cdots, \omega_{20}\right)'$. These estimates are then transformed to obtain the parameters of interest $\boldsymbol{\theta} = \left(\boldsymbol{\mu}_t^{*\prime}, \boldsymbol{\rho}_t{}', \boldsymbol{\omega}^{2\prime}\right)'$ and standard errors for $\boldsymbol{\theta}$ are obtained by the multivariate delta method using

$$\mathrm{Acov}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{N}\mathbf{CH}\boldsymbol{\Xi}\mathbf{H}'\mathbf{C}' \qquad \mathbf{C} = \frac{\partial \boldsymbol{\theta}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}'} \tag{22}$$

Previous attempts to estimate these models by minimizing a function of $\boldsymbol{\theta}$ directly resulted in a wild variability of parameter estimates and many non-convergent solutions. Sample sizes larger than 10,000 observations were needed to obtain results similar to those obtained here with 500 observations.

4. Thurstonian Correlation Structure Models for Multiple Judgment Paired Comparison Data

In the previous sections we considered two classes of Thurstonian models for paired comparisons data. In both cases, the pattern probabilities are given by (12) and (13). The models differ in the restrictions imposed on the thresholds and tetrachoric correlations:

(a) Thurstone-Takane class with equal pair specific errors, where $\boldsymbol{\tau} = -\mathbf{DA}\boldsymbol{\mu}_t$, and $\boldsymbol{P}_{z^*} = \mathbf{D}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \omega^2\mathbf{I}\right)\mathbf{D}$.

(b) Thurstone-Takane class with unequal pair specific errors, where $\boldsymbol{\tau} = -\mathbf{DA}\boldsymbol{\mu}_t$, and $\boldsymbol{P}_{z^*} = \mathbf{D}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \boldsymbol{\Omega}^2\right)\mathbf{D}$.

In this section we shall introduce a third class of Thurstonian models for paired comparisons data with pattern probabilities also given by (12) and (13):

(c) Thurstone correlation structure model, where $\boldsymbol{\tau} = -\mathbf{A}\boldsymbol{\mu}_t$, and $\boldsymbol{P}_{z^*} = \text{Off}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'\right)$, where Off(•) is used to denote the restrictions imposed on the off-diagonal elements of a correlation matrix.

The rationale for this third class of models is as follows: The Thurstone and Thurstone-Takane models are mean and covariance structure models in the sense that they impose constraints on the mean vector and covariance matrix of $\mathbf{y}^*$, the pairwise differences among the unobserved continuous preferences $\mathbf{t}$. However, as only the binary choices are observed, the covariance matrix of $\mathbf{y}^*$ can only be estimated from a correlation matrix. Moreover, to estimate the parameters of these models one must resort to pre and post-multiply the covariance structure by the inverse of a diagonal matrix of model-based standard deviations. This results in complex non-linear restrictions on the thresholds and tetrachoric correlations.

The class of models (c) is a solution –different from Takane's (1987)- to the problem of specifying a Thurstonian model for multiple judgment pair comparison data that assigns non-zero probabilities to all paired comparisons patterns. It amounts to specifying a mean and correlation structure model for $\mathbf{y}^*$ consistent with Thurstone's (1927) seminal ideas rather than a mean and covariance structure model. Correlation structure models are commonly used in situations where modeling the variances is of no interest (McDonald, 1975). Within the present framework a mean and correlation structure model imposes restrictions on the means and correlations of $\mathbf{y}^*$, yet leaves the structure for the variances of $\mathbf{y}^*$ unspecified.

Now, from (6) the mean and covariance structures for $\mathbf{y}^*$ implied by Thurstone's model are $\boldsymbol{\mu}_{y^*} = \mathbf{A}\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_{y^*} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'$. Thus, we shall assume $\boldsymbol{\mu}_{y^*} = \mathbf{A}\boldsymbol{\mu}_t$. Also, by parsimony we shall assume $\boldsymbol{P}_{y^*} = \text{Off}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'\right)$. That is, we assume that the restrictions on the correlations among $\mathbf{y}^*$ have the same functional form as the restrictions of Thurstone's model on the covariances among $\mathbf{y}^*$. Then, the probability of any paired comparisons pattern under this Thurstonian mean and correlation structure model is

$$\Pr\left(\bigcap_{l=1}^{\tilde{n}} y_l\right) = \int \underset{\mathbf{R}}{\cdots} \int \phi_{\tilde{n}}\left(\mathbf{y}^* : \boldsymbol{\mu}_{y^*}, \boldsymbol{P}_{y^*}\right) d\mathbf{y}^* \tag{23}$$

with limits of integration (7). These pattern probabilities are unchanged when we transform $\mathbf{y}^*$ using $\mathbf{z}^* = \mathbf{y}^* - \boldsymbol{\mu}_{y^*}$, where at $y_l^* = 0$, $\tau_l := -\mu_l^*$. As a result, $\boldsymbol{\mu}_{z^*} = \mathbf{0}$ and

$$\boldsymbol{\tau} = -\mathbf{A}\boldsymbol{\mu}_t \qquad \boldsymbol{P}_{z^*} = \text{Off}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'\right) \tag{24}$$

where now the pattern probabilities are given by (12), (13) and (24). Again, to identify an unrestricted model, it is necessary to set the location of the elements of $\boldsymbol{\mu}_t$ and the location of the elements in each of the rows (columns) of $\boldsymbol{\Sigma}_t$. Arbitrarily, we shall set $\mu_n = 0$ and $\boldsymbol{\Sigma}_t = \boldsymbol{P}_t$.

We immediately see that our proposed Thurstone correlation structure model has the same pattern probabilities as the Thurstone-Takane model with Maydeu-Olivares (2001) identification solution $\boldsymbol{\Omega}^2 = \mathbf{I} - \text{Diag}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'\right)$. Our Thurstone correlation structure model and Maydeu-Olivares' model are statistically equivalent, yet substantively different. The correlation structure model makes no structural assumptions about the variances of $\mathbf{y}^*$, no pair specific random errors are assumed, and hence no improper solutions can be obtained. However, as these two models are statistically equivalent, the simulation results reported by Maydeu-Olivares (2001) for his model are applicable to our correlation structure model as well.

An unrestricted correlation structure model is more restricted than an unrestricted Thurstone-Takane model assuming $\boldsymbol{\Omega}^2$ diagonal as the latter has $(\tilde{n} - 1)$ additional parameters. Furthermore, although the Thurstonian correlation structure model has the same number of identified parameters as the Thurstone-Takane model when $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$, these models are not equivalent. Also, the Thurstone-Takane model with $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ and $\boldsymbol{\Sigma}_t$ diagonal is not equivalent to the Thurstonian correlation structure model with $\boldsymbol{\Sigma}_t$ diagonal. Nonetheless, the Thurstone-Takane model with $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ and $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$ is equivalent to the Thurstonian correlation structure model with $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$. In this special case, denoting by $\tilde{\mu}_i$ and $\tilde{\sigma}^2$ the parameters of Thurstone-Takane model and μ_i and σ^2 the parameters of the Thurstone correlation structure model we have

$$\tilde{\mu}_i = \mu_i\sqrt{\frac{1}{1-2\sigma^2}} \qquad \tilde{\sigma}^2 = \frac{\sigma^2}{1-2\sigma^2}. \tag{25}$$

The choice between employing a covariance or a correlation structure within a Thurstonian framework, should be grounded substantively. If one believes that the variances of $\mathbf{y}^*$ ought to be modeled and one feels at ease with Takane's (1987) proposal of adding a vector of pair specific errors to Thurstone then a Thurstone-Takane model should be employed. On the other hand, if one believes that the variances of $\mathbf{y}^*$ should not be modeled then a correlation structure model should be employed. Statistically, large sample sizes are needed to accurately estimate a Thurstone-Takane model with $\boldsymbol{\Omega}^2$ diagonal. Rather small sample sizes are needed to estimate a Thurstone-Takane model with $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$. However, even smaller sample sizes are needed to estimate our proposed Thurstonian correlation model. This can be seen in Table 1. The parameter estimates obtained by Maydeu-Olivares (2001) are more accurate and less variable than those obtained for the Thurstone-Takane model with $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$. Thus, if only a very small sample is available to estimate a very large model, on statistical grounds, one should employ a correlation structure model. In addition, special purpose software is needed to estimate covariance structure models, whereas commercially available software such as MPLUS (Muthén & Muthén, 1998) can be employed to estimate correlation structure models. To illustrate this point, in Appendix 1 we provide code to estimate Thurstonian correlation structure models using MPLUS for the first of the applications described in the next section.

5. Applications

We now provide three different applications in which we fit: (a) an unrestricted Thurstonian covariance structure model assuming that $\boldsymbol{\Omega}^2$ is a diagonal matrix, (b) an unrestricted Thurstonian covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$, and (c) an unrestricted Thurstonian correlation structure model. Our aims are to compare empirically models (b) and (c) which have the same number of parameters but are not equivalent, and to investigate empirically if relaxing the assumption $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ letting $\boldsymbol{\Omega}^2$ be a diagonal matrix improves the fit of the covariance structure model.

5.1 Modeling preferences for compact cars

Maydeu-Olivares (2001) reported a study in which preferences for compact cars among college students where investigated. The sample size was 289 and he published a subset of the data consisting of the responses to these four cars {Opel Corsa, Renault Clio, Seat Ibiza, Volkswagen Polo}.

The structural and overall goodness of fit tests for the three models under consideration are provided in Table 3. As can be seen in this table, the differences in overall fit between the unrestricted covariance structure model with $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ and the unrestricted correlation structure are negligible. Also, there is no improvement in the fit of the covariance structure model when the assumption $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ is relaxed by assuming that $\boldsymbol{\Omega}^2$ is diagonal.

Insert Tables 3 and 4 about here

Regarding the parameter estimates, it is important to realize that when it is assumed that $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ the free parameters are estimated relative to the value assigned to ω^2. Similarly, when it is assumed that $\boldsymbol{\Omega}^2$ is a diagonal matrix, the free parameters are estimated relative to the value assigned to $\omega^2_{\tilde{n}}$. In our discussion of the identification of these models, we suggested –for ease of implementation- assigning a value of one to ω^2 and $\omega^2_{\tilde{n}}$. However, arbitrarily choosing another value for ω^2 and $\omega^2_{\tilde{n}}$ results in a different set of parameter estimates with the same pattern probabilities. This indeterminacy is unsatisfactory from an applied perspective as one is interested in meaningfully interpreting the parameter estimates. In Appendix 2 we describe a solution to this problem consisting in a reparameterization of the model that yields parameter estimates for all the free parameters and ω^2 when $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ is assumed, or for all the free parameters and $\omega^2_{\tilde{n}}$ when $\boldsymbol{\Omega}^2$ is assumed to be diagonal. Although this reparameterization solution is more complex to implement, we strongly recommend its use in applications to overcome the indeterminacy just described.

In Table 4 we provide the parameter estimates and standard errors for the covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ obtained using this reparameterization. The value estimated for ω^2 was 0.13 with a standard error of 0.04. With these data, fixing $\omega^2 = 1$ to identify the model causes the parameter estimates in $\boldsymbol{P}_t$ to lie outside the range [-1, 1]. A fixed value of 1 for ω^2 is too large in this application. In Table 4 we also provide the parameter estimates and standard errors for the unrestricted correlation structure model. We notice that the same substantive conclusions are obtained using both approaches. However, the standard errors for the elements of $\boldsymbol{P}_t$ in the correlation structure model are considerably smaller, and because this is a linear model, they are more homogeneous.

Given the results observed in Table 4, we conjectured that a model assuming $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$ could adequately fit these data. Thus, we fitted an unrestricted Thurstonian correlation structure model assuming $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$, and a Thurstone-Takane model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ and $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$ to these data. These two models are equivalent. The goodness of fit tests are also reported in Table 3. It is unclear how to perform nested tests using the mean and variance corrected statistics. However, Satorra and Bentler (2001) have recently provided expressions for performing nested tests using the mean corrected statistics. Using $\alpha = 0.01$, nested tests of the overall restrictions suggests that no fit improvement is obtained by assuming either an unrestricted covariance or an unrestricted correlation structure model over a model with the restriction $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$. Thus, the most parsimonious representation for these data is a model with the restriction $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$. With standard errors in parentheses, the parameter estimates in the correlation structure metric are $\mu_1 = 0.20$ (0.07), $\mu_2 = -0.16$ (0.07), $\mu_3 = -0.11$ (0.07), $\sigma^2 = 0.45$ (0.01), and in the covariance structure metric are $\mu_1 = 0.65$ (0.23), $\mu_2 = -0.50$ (0.23), $\mu_3 = -0.36$ (0.22), $\sigma^2 = 4.78$ (1.11), where we fixed $\omega^2 = 1$.

5.2 Modeling paired comparisons in a personality assessment task

In personnel selection tasks it is often necessary to obtain some personality assessment of the applicants. Rating scales are ill-suited to this purpose as applicants naturally distort their responses to match a desirable personality profile. This desirability effect can be greatly reduced by using a paired comparisons design in which respondents are forced to choose between pairs of desirable stimuli. At the Catalan Police Academy (Spain) a personality assessment is performed on all trainees using a paired comparisons design. Given a set of adjectives describing personality characteristics, trainees are presented with all adjective pairs and are asked to choose the adjective that best describes their personality within each pair. In Table 5 we provide all binary response patterns and observed frequencies yielded by all male trainees (580) assessed during 2001 from one particular police corps to four of the adjectives {competent, orderly, reliable, and resolved}. These data will be analyzed in this example.

Insert Tables 5 to 8 about here

The overall goodness of fit tests for the three models under consideration are presented in Table 6. As can be seen in this table, the covariance structure model with unequal paired specific error variances provides a good fit to these data. The other two models fit rather poorly while the covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ fits better than a correlation structure model.

In Table 7 we provide the parameter estimates and standard errors for the good-fitting model. These were estimated by fixing $\omega_6^2 = 1$. The reparameterization approach described in Appendix 2 yields a value of 0.99 for ω_6^2 (SE = 0.37). Hence the fixed value assigned to ω_6^2 is appropriate. As can be seen in this table, the parameter estimates for $\boldsymbol{\Omega}^2$ differ considerably, ranging from 0.25 (SE = 0.26) to 4.45 (SE = 1.83). Thus, it is not surprising that a model assuming equal pair specific error variances fits poorly. Interestingly, the ordering of the mean preferences is {orderly, resolved, competent, and reliable}. Thus, for instance, police trainees prefer to describe themselves as orderly rather than resolved. Also, it is interesting to see that there are only two significant parameter estimates in $\boldsymbol{P}_t$: Police trainees describing themselves as competent are more likely to describe themselves as reliable, and police trainees describing themselves as competent are more likely to describe themselves as resolved.

5.3 Modeling preferences for celebrities

Kroeger (1992) replicated a classical experiment by Rumelhart and Greeno (1971) in which respondents were presented with pairs of celebrities and they were asked to select the celebrity with whom they would rather spend an hour of conversation. Here we shall analyze a subset of Kroeger's data consisting of the females' responses (96 subjects) to the paired comparisons involving the set of former U.S. first ladies (Barbara Bush, Nancy Reagan, and Hillary Clinton) and athletes (Bonnie Blair, Jackee Joyner-Kersee, and Jeniffer Capriati).

Insert Tables 8 and 9 about here

Again, we fitted the three models under consideration to these 15 paired comparisons. The overall goodness of fit tests for these models are presented in Table 8. As can be seen in this table, the correlation structure model does not fit these data well. But the covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ provides a good fit to these data. Furthermore, no improvement in fit is apparent by relaxing the restriction $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ and assuming that $\boldsymbol{\Omega}^2$ is a diagonal matrix.

In Table 9 we provide the parameter estimates and standard errors for the covariance structure model with $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ using the reparameterization approach of Appendix 2. The estimate of ω^2 is 0.26. As can be seen in this table a preference for Barbara Bush is strongly related to a preference for Nancy Reagan. Furthermore, preferences for the three athletes are all positively correlated. These are the only significant correlations in $\boldsymbol{P}_t$ given the small sample size available, but the non-significant correlations also seem to agree with what one would expect. For instance, preferences for Hillary Clinton are negatively related to preferences for Barbara Bush and Nancy Reagan. Finally, the ordering of the mean preferences for these celebrities under this model is {Hillary Clinton, Jackie Joyner- Kersee, Barbara Bush, Jeniffer Capriati, Bonnie Blair, and Nancy Reagan} although the large standard errors obtained with this small sample size leads to numerous ties in this ordering.

6. Discussion and conclusions

In this paper we have considered mean and covariance structure and mean and correlation structure Thurstonian models for multiple judgment paired comparison data. Letting $\mathbf{y}^*$ denote the vector of pairwise differences among the unobserved continuous preferences assumed to underlie the binary choices in Thurstonian models, in both cases, the mean structure is $\boldsymbol{\mu}_{y^*} = \mathbf{A}\boldsymbol{\mu}_t$. In the covariance structure model it is assumed that $\boldsymbol{\Sigma}_{y^*} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \boldsymbol{\Omega}^2$, and in the correlation structure model it is assumed that $\boldsymbol{P}_{y^*} = \text{Off}\left(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'\right)$. These correlation and covariance structures are not equivalent models because the covariance structure model is not scale invariant. Furthermore, the correlation structure model is more restrictive than the covariance structure model. In Maydeu-Olivares (2001) this issue was overlooked and he proposed identification restrictions for the covariance structure model that were unnecessarily restrictive.

We believe that the choice between employing a covariance or a correlation structure within a Thurstonian framework should be motivated substantively. However, it is also important to consider model fit and estimation issues in choosing between these models. Model fit issues are important because substantive conclusions drawn from a poor fitting model are meaningless. Estimation issues are also important since in many applications of multiple judgment paired comparisons data only small samples are collected and the contingency tables can be extraordinarily sparse. We have investigated in some detail model fit issues in applications and estimation issues in simulation studies. We have considered three unrestricted Thurstonian models: (a) a covariance structure model assuming that $\boldsymbol{\Omega}^2$ is a diagonal matrix, (b) a covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$, and (c) a correlation structure model. Models (b) and (c) have the same number of identified parameters yet they are not equivalent models. Model (a) is more general.

We have reported here three numerical examples that illustrate well our model-fit findings. In not any of the applications we have investigated, the correlation structure model fitted the data better than the covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$, although the differences in fit were in some cases small. This was the case in the first of the applications we reported. Regarding the assumptions on $\boldsymbol{\Omega}^2$ in covariance structure models, we have found that in many applications no appreciable fit improvement is obtained by assuming that $\boldsymbol{\Omega}^2$ is a diagonal matrix rather than $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$. This was the case in the third application we reported. However, this statement must be qualified, as most applications we have considered consisted of small samples. When larger samples are available, a substantial fit improvement assuming $\boldsymbol{\Omega}^2$ diagonal over $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ is more likely to be found. This was the case in the second application we reported.

The simulation studies performed suggest that samples as small as 100 observations suffice to estimate covariance structure models for 7 stimuli assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$. Larger sample sizes are needed to estimate covariance structure models when $\boldsymbol{\Omega}^2$ is assumed to be a diagonal matrix. Correlation structure models can be estimated with smaller sample sizes than covariance structure models assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$.

Taking together model fit and estimation results we tentatively conclude that from a statistical viewpoint the use of a covariance structure model with $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ is to be recommended. If a large sample is available, we conjecture that a covariance structure model assuming that $\boldsymbol{\Omega}^2$ is a diagonal matrix may yield a substantial fit improvement. However, if only a extremely small sample is available relative to the size of the model, a correlation structure model should be considered, as this model can be estimated with very small samples.

References

David, H.A. (1988). The method of paired comparisons. London: Griffin.

Kroeger, K. (1992). Unpublished dataset. University of Illinois, Urbana-Champaign.

McDonald, R.P. (1975). Testing pattern hypotheses for correlation matrices. Psychometrika, 40, 253-255.

Maydeu-Olivares, A. (1999). Thurstonian modeling of ranking data via mean and covariance structure analysis. Psychometrika, 64, 325-340.

Maydeu-Olivares, A. (2001). Limited information estimation and testing of Thurstonian models for paired comparison data under multiple judgment sampling. Psychometrika, 66, 209-228.

Muthén, B. (1978). Contributions to factor analysis of dichotomous variables. Psychometrika, 43, 551-560.

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

Muthén, B. (1993). Goodness of fit with categorical and other non normal variables. In K.A. Bollen & J.S. Long [Eds.] Testing structural equation models (pp. 205-234). Newbury Park, CA: Sage.

Muthén, B., du Toit, S.H.C. & Spisic, D. (in press). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Psychometrika.

Muthén, L. & Muthén, B. (1998). MPLUS. Los Angeles, CA: Muthén & Muthén.

Rummelhart, D.L. & Greeno, J.G. (1971). Similarity between stimuli: An experimental test of the Luce and Restle choice models. Journal of Mathematical Psychology, 8, 370-381.

Satorra, A. & Bentler, P.M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye and C.C. Clogg (Eds.), Latent variable analysis: Applications to developmental research (pp. 399-419). Thousand Oaks, CA: Sage.

Satorra, A. & Bentler, P.M. (2001). A scaled difference chi-square test statistic for moment structure analysis. Psychometrika, 66,

Takane, Y. (1987). Analysis of covariance structures and probabilistic binary choice data. Communication and Cognition, 20, 45-62.

Thurstone, L.L. (1927). A law of comparative judgment. Psychological Review, 79, 281-299.

TABLE 1

Simulation results for an unrestricted Thurtone-Takane model with 7 stimuli: Parameter estimates and standard errors

	$\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$						$\boldsymbol{\Omega}^2$ *diagonal*						*Maydeu-Olivares (2001)*		
	N = 100			N = 500			N = 500			N = 1000			N = 100		
par	$\bar{x}$ est.	$\bar{x}$ SEs	std est.	$\bar{x}$ est.	$\bar{x}$ SEs	std est.	$\bar{x}$ est.	$\bar{x}$ SEs	std est.	$\bar{x}$ est.	$\bar{x}$ SEs	std est.	$\bar{x}$ est.	$\bar{x}$ SEs	std est.
μ = 0.5	0.50	0.11	0.11	0.50	0.05	0.05	0.50	0.11	0.12	0.50	0.08	0.08	0.50	0.09	0.09
μ = 0	0.00	0.10	0.10	0.00	0.05	0.05	0.00	0.05	0.05	0.00	0.04	0.04	0.00	0.09	0.09
μ = -0.5	-0.51	0.12	0.13	-0.50	0.05	0.06	-0.51	0.09	0.10	-0.50	0.07	0.07	-0.51	0.11	0.11
ρ = 0.8	0.79	0.08	0.08	0.80	0.04	0.04	0.79	0.09	0.10	0.80	0.06	0.06	0.80	0.05	0.05
ρ = 0.7	0.69	0.10	0.10	0.70	0.04	0.04	0.68	0.12	0.14	0.70	0.08	0.08	0.70	0.06	0.06
ρ = 0.6	0.59	0.12	0.12	0.60	0.05	0.05	0.58	0.15	0.17	0.60	0.10	0.11	0.60	0.06	0.07
ω^2 = 1	--	--	--	--	--	--	1.13	0.58	0.69	1.06	0.38	0.41	--	--	--

Notes: ULS estimation; 1000 replications;

TABLE 2

Simulation results for an unrestricted Thurtone-Takane model with 7 stimuli: Goodness of fit tests

						Nominal rates										
Null hypothesis	*Model*	*N*	*Stat.*	*Mean*	*Var.*	*1%*	*5%*	*10%*	*20%*	*30%*	*40%*	*50%*	*60%*	*70%*	*80%*	*90%*
$H_0 : \boldsymbol{\kappa} = \boldsymbol{\kappa}(\boldsymbol{\theta})$	$\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ [a]	100	T_s	210.9	539.4	3.2	11.2	20.4	32.3	42.5	54.1	63.1	70.8	78.7	84.2	91.2
			T_a	56.6	37.3	0.1	0.6	2.8	13.5	31.1	47.3	65.8	79.7	89.2	95.8	99.7
		500	T_s	204.6	472.1	1.5	6.1	12.3	23.5	32.8	41.5	49.8	59.1	69.4	79.5	89.1
			T_a	122.5	162.5	0.4	2.4	6.0	16.7	29.3	39.8	50.9	62.7	75.8	86.2	94.3
	$\boldsymbol{\Omega}^2$ *diagonal* [b]	500	T_s	183.8	421.2	1.4	5.5	11.2	22.4	31.4	39.4	47.7	57.7	68.5	77.6	89.4
			T_a	114.1	156.7	0.4	2.2	6.0	17.2	28.0	37.7	48.4	61.0	73.6	84.8	93.4
		1000	$\tilde{T}_s$	184.0	439.1	1.9	6.5	12.2	21.0	30.3	40.6	48.4	57.9	67.6	77.7	88.9
			$\tilde{T}_a$	132.8	223.8	0.9	4.2	8.6	17.9	27.4	39.8	48.8	60.6	72.0	81.7	92.7
$H_0 : \boldsymbol{\pi} = \boldsymbol{\pi}(\boldsymbol{\theta})$	$\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ [a]	100	$\tilde{T}_s$	206.1	2894.1	17.7	24.9	29.5	35.7	39.9	43.6	47.3	51.2	55.6	60.0	66.4
			$\tilde{T}_a$	21.5	30.9	0.3	3.4	7.8	17.7	28.5	40.1	51.8	63.4	74.3	84.7	95.4
		500	$\tilde{T}_s$	205.8	3148.3	18.1	25.2	28.8	36.0	39.2	43.6	46.7	50.2	54.3	58.9	63.8
			$\tilde{T}_a$	26.6	52.4	1.1	5.0	10.6	19.8	29.9	39.7	50.0	59.9	69.5	81.5	91.8
	$\boldsymbol{\Omega}^2$ *diagonal* [b]	500	$\tilde{T}_s$	185.1	2668.2	16.3	23.0	27.2	33.9	37.5	42.4	46.4	50.6	54.0	59.6	66.1
			$\tilde{T}_a$	24.6	47.7	1.2	5.6	10.0	17.7	28.2	38.6	50.2	60.7	70.7	81.3	91.8
		1000	$\tilde{T}_s$	185.3	2498.1	16.8	24.1	28.6	34.6	38.6	43.0	46.8	50.5	54.1	58.7	65.6
			$\tilde{T}_a$	25.7	48.3	1.4	4.7	9.5	19.4	29.9	39.9	39.6	50.0	59.9	70.2	81.8

Notes: ULS estimation; 1000 replications; [a] 204 d.f.; [b] 184 d.f.

TABLE 3

Goodness of fit tests for the compact cars data

	H_0: $\boldsymbol{\kappa} = \boldsymbol{\kappa}(\boldsymbol{\theta})$				H_0: $\boldsymbol{\pi} = \boldsymbol{\pi}(\boldsymbol{\theta})$			
Model	*stat*	*val.*	*df*	*p-val*	*stat*	*val.*	*df*	*p-val*
cov. structure	T	8.51	--	--	$\tilde{T}$	2.21	--	--
unrestricted	T_s	7.98	7	0.33	$\tilde{T}_s$	24.22	7	<0.01
$\boldsymbol{\Omega}^2$ diagonal	T_a	6.65	5.84	0.34	$\tilde{T}_a$	20.20	4.10	<0.01
cov. structure	T	10.11	--	--	$\tilde{T}$	1.91	--	--
unrestricted	T_s	12.93	12	0.37	$\tilde{T}_s$	7.35	12	0.01
$\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$	T_a	8.77	8.13	0.37	$\tilde{T}_a$	18.54	4.40	0.02
corr. structure	T	9.21	--	--	$\tilde{T}$	2.18	--	--
unrestricted	T_s	10.13	12	0.60	$\tilde{T}_s$	28.77	12	<0.01
	T_a	7.82	9.27	0.58	$\tilde{T}_a$	22.22	4.57	0.01
corr. structure	T	55.91	--	--	$\tilde{T}$	3.87	--	--
$\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$	T_s	30.46	17	0.02	$\tilde{T}_s$	44.44	17	<0.01
	T_a	15.69	8.75	0.07	$\tilde{T}_a$	22.87	7.71	0.01

Notes: T_s and $\tilde{T}_s$ are Satorra-Bentler's scaled statistics, T_a and $\tilde{T}_a$ are Satorra-Bentler's adjusted statistics. The correlation structure model with $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$ is equivalent to a covariance structure model with $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$ and $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$.

TABLE 4

Estimated parameter estimates and standard errors for an unrestricted model applied to the cars data

	$\boldsymbol{P}_t$	Opel Corsa	Seat Ibiza	VW Polo	Renault Clio	$\boldsymbol{\mu}_t$	
covariance structure	Opel Corsa	1	0.66 (0.04)	0.50 (0.04)	0.56 (0.04)	0.20 (0.07)	correlation structure
	Seat Ibiza	0.54 (0.10)	1	0.56 (0.04)	0.50 (0.04)	-0.15 (0.07)	
	VW Polo	0.27 (0.08)	0.35 (0.06)	1	0.50 (0.04)	-0.11 (0.07)	
	Renault Clio	0.35 (0.13)	0.27 (0.14)	0.24 (0.09)	1	0 *(fixed)*	
	$\boldsymbol{\mu}_t$	0.22 (0.08)	-0.18 (0.09)	-0.14 (0.09)	0 *(fixed)*		

<u>Notes</u>: $N = 289$; standard errors in parentheses; the covariance structure model assumes $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$; $\hat{\omega}^2 = 0.13$ (0.04).

TABLE 5

Observed frequencies of paired comparisons patterns in the personality profile data

no.	pattern	obs.	no.	pattern	obs.	no.	pattern	obs.	no.	pattern	obs.
* 1	111111	1	17	101111	1	*33	011111	43	*49	001111	19
* 2	111110	14	18	101110	2	*34	011110	81	50	001110	3
3	111101	2	19	101101	0	35	011101	2	51	001101	0
* 4	111100	25	20	101100	0	36	011100	12	52	001100	1
* 5	111011	1	*21	101011	1	37	011011	7	*53	001011	7
6	111010	4	22	101010	0	38	011010	19	54	001010	1
* 7	111001	0	*23	101001	2	39	011001	1	55	001001	3
* 8	111000	5	24	101000	2	40	011000	10	56	001000	0
9	110111	0	25	100111	0	41	010111	11	*57	000111	13
10	110110	10	26	100110	3	*42	010110	78	*58	000110	23
11	110101	3	27	100101	1	43	010101	3	59	000101	2
*12	110100	33	28	100100	3	*44	010100	62	*60	000100	8
13	110011	5	29	100011	0	45	010011	0	*61	000011	4
14	110010	5	30	100010	1	46	010010	9	62	000010	0
15	110001	1	*31	100001	2	47	010001	3	*63	000001	2
*16	110000	12	*32	100000	2	48	010000	7	*64	000000	5

Notes: $N = 580$; * transitive patterns

TABLE 6

Overall goodness of fit tests for the unrestricted models applied to the personality assessment data

Model	*stat*	*val.*	*df*	*p-val*
cov. structure $\boldsymbol{\Omega}^2$ diagonal	$\tilde{T}$	0.65	--	--
	$\tilde{T}_s$	8.75	7	0.27
	$\tilde{T}_a$	7.98	2.72	0.28
cov. structure $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$	$\tilde{T}$	19.85	--	--
	$\tilde{T}_s$	201.43	12	<0.01
	$\tilde{T}_a$	169.21	4.16	<0.01
corr. structure	$\tilde{T}$	26.74	--	--
	$\tilde{T}_s$	261.28	12	<0.01
	$\tilde{T}_a$	220.83	4.24	<0.01

Notes: $N = 580$; $\tilde{T}_s$ and $\tilde{T}_a$ are the scaled for mean and adjusted for mean and variance statistics.

TABLE 7

Estimated parameter estimates and standard errors for an unrestricted covariance structure model assuming $\boldsymbol{\Omega}^2$ diagonal applied to the personality assessment data

$\boldsymbol{P}_t$	competent	orderly	reliable	resolved
competent	1			
orderly	0.48 (0.27)	1		
reliable	**0.44** (0.20)	0.25 (0.31)	1	
resolved	**0.60** (0.22)	0 (0.50)	0.10 (0.37)	1
$\boldsymbol{\mu}_t$	-0.11 (0.07)	0.68 (0.21)	-1.24 (0.20)	0 *(fixed)*

ω_1^2	ω_2^2	ω_3^2	ω_4^2	ω_5^2	ω_6^2
0.25 (0.26)	0.59 (0.31)	0.80 (0.78)	4.45 (1.83)	1.39 (0.84)	1 *(fixed)*

Notes: $N = 580$; standard errors in parentheses; correlations in bold are significant at $\alpha = 0.05$.

TABLE 8

Overall goodness of fit tests for the unrestricted models applied to the celebrities data

Model	*stat*	*val.*	*df*	*p-val*
cov. structure $\boldsymbol{\Omega}^2$ diagonal	$\tilde{T}$	9.94	--	--
	$\tilde{T}_s$	135.05	86	<0.01
	$\tilde{T}_a$	49.44	15.54	0.02
cov. structure $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$	$\tilde{T}$	8.63	--	--
	$\tilde{T}_s$	124.73	100	0.05
	$\tilde{T}_a$	41.00	16.09	0.16
corr. structure	$\tilde{T}$	12.92	--	--
	$\tilde{T}_s$	172.75	100	<0.01
	$\tilde{T}_a$	57.82	17.05	<0.01

Notes: $N = 96$; $\tilde{T}_s$ and $\tilde{T}_a$ are the scaled for mean and adjusted for mean and variance statistics.

TABLE 9

Estimated parameter estimates and standard errors for an unrestricted covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ applied to the celebrities data

$\boldsymbol{P}_t$	Barbara Bush	Nancy Reagan	Hillary Clinton	Bonnie Blair	Jackie Joyneer	Jennifer Capriati
Barbara Bush	1					
Nancy Reagan	**0.74** (0.12)	1				
Hillary Clinton	-0.24 (0.49)	-0.39 (0.54)	1			
Bonnie Blair	0.23 (0.27)	0.34 (0.25)	-0.21 (0.45)	1		
Jackie Joyneer	-0.03 (0.25)	0.03 (0.24)	-0.49 (0.47)	**0.37** (0.08)	1	
Jennifer Capriati	0.25 (0.23)	0.20 (0.26)	-0.58 (0.55)	**0.43** (0.20)	**0.31** (0.10)	1
$\boldsymbol{\mu}_t$	0.09 (0.15)	-0.28 (0.16)	0.44 (0.21)	-0.27 (0.14)	0.16 (0.14)	0 *(fixed)*

Notes: $N = 96$; standard errors in parentheses; correlations in bold are significant at $\alpha = 0.05$; $\hat{\omega}^2 = 0.26$ (0.09).

APPENDIX 1

Estimation of Thurstonian correlation structure models using MPLUS

Let $\mathbf{y}^*$ be a $\tilde{n}$-multidimensional normal density categorized according to a set of thresholds. When there are no exogenous variables, the measurement model used in MPLUS (Muthén & Muthén, 1998) is

$$\mathbf{y}^* = \boldsymbol{\nu} + \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon} \tag{26}$$

where $\boldsymbol{\eta}$ denotes a p-dimensional vector of latent variables, $\boldsymbol{\varepsilon}$ denotes a $\tilde{n}$-dimensional vector of residual measurement errors, $\boldsymbol{\nu}$ denotes a ñ-dimensional vector of measurement intercepts and $\boldsymbol{\Lambda}$ is a $\tilde{n}$ x p matrix of measurement slopes (factor loadings). We further assume that

$$\begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta} \end{pmatrix}\right) \tag{27}$$

MPLUS implements a three-stage estimation approach equivalent to the approach described here except that the asymptotic covariance matrix of $\sqrt{N}\left(\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}\right)$, $\boldsymbol{\Xi}$, is computed as in Muthén (1984) rather than as in Muthén (1978) as we do here. In the third stage, the user may choose ULS, DWLS or WLS estimation.

Since when ULS is selected, no standard errors nor goodness of fit tests are provided, we shall provide code here to estimate an unrestricted Thurstonian correlation structure model using DWLS to the compact cars data. In this code we assume that the individual data is provided, a file of 289 observations consisting of 6 columns separated by spaces. To estimate this model, we let $\boldsymbol{\nu} = \mathbf{0}, \boldsymbol{\Lambda} = \mathbf{A}, \boldsymbol{\alpha} = \boldsymbol{\mu}_t$, and $\boldsymbol{\Psi} = \boldsymbol{P}_t$. That is, we simply estimate a factor model with n latent variables, a fixed matrix of factor loadings $\mathbf{A}$, and we estimate the factor means and inter-factor correlations. The matrix $\boldsymbol{\Theta}$ is not estimated in MPLUS. To identify the model we fix the last factor mean, and we fix all the factor variances to 1. Finally, we need to fix the MPLUS thresholds to zero.

```
TITLE: DWLS estimation of Maydeu-Olivares (2001) data
DATA: FILE IS pc.dat;
VARIABLE: NAMES ARE y1-y6;
      CATEGORICAL = y1-y6;
ANALYSIS:  TYPE = MEANSTRUCTURE;
           ESTIMATOR = WLSM ;
!     mean corrected test statistic
!     WLSMV yields the mean and variance corrected statistic
MODEL:
f1 BY y1@1 y2@1 y3@1;
f2 by y1@-1 y4@1 y5@1;
f3 by y2@-1 y4@-1 y6@1;
f4 by y3@-1 y5@-1 y6@-1;
```

```
!     fixed factor loadings, this is the A matrix
f1-f4@1;                                                    (28)
!     factor variances are fixed at 1
!     factor correlations are free parameters
[y1$1-y6$1@0];
!     thresholds fixed to zero, by default they are free
[f1 f2 f3 ];
!     factor means free, default is fixed to zero
!     measurement intercepts are zero by default
OUTPUT: TECH1;
```

MPLUS only provides tests of the structural restrictions implied by the model. If both the mean (T_s) and the mean and variance (T_a) tests statistics are of interest, the model must be run twice. Despite the difference of estimation method (DWLS vs. ULS) and formula used to estimate $\boldsymbol{\Xi}$, the results obtained using MPLUS closely match those reported in Table 3: $T_s =$ 10.69, p =0.56, and $T_a = 8.01$, $p = 0.53$.

Using straightforward modifications on this code one can estimate a variety of restricted Thurstonian correlation structure models. For instance, using

```
f4 with f1-f3@0;
f3 with f1-f2@0;
f2 with f1@0;
```

in place of (28) yields a model where $\boldsymbol{\Sigma}_t$ is a diagonal matrix (Thurstone's Case III model). Alternatively, using

```
f1-f4 (1);
f4 with f1-f3@0;
f3 with f1-f2@0;
f2 with f1@0;
```

in place of (28) yields a model with the restriction $\boldsymbol{\Sigma}_t = \sigma^2\mathbf{I}$ (Thurstone's Case V model).

APPENDIX 2

A reparameterization approach to estimate Thurstonian covariance structure models

Consider an unrestricted covariance structure model assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$, so that $\boldsymbol{\tau} = -\mathbf{DA}\boldsymbol{\mu}_t$, and $\boldsymbol{P}_{z^*} = \mathbf{D}\left(\mathbf{A}\boldsymbol{P}_t\mathbf{A}' + \omega^2\mathbf{I}\right)\mathbf{D}$. This model can be identified by letting $\mu_n = 0$, and $\omega^2 = 1$. However, arbitrarily choosing another value for ω^2 results in a different set of parameter estimates with the same pattern probabilities. From an applied perspective, this indeterminacy is troublesome as one is interested in meaningfully interpreting the parameter estimates. Here we describe a reparameterization approach that yields an estimate of ω^2 as well, thus overcoming this indeterminacy.

Maydeu-Olivares (1999: Appendix 2) pointed out that because **A** is of rank n - 1, it can be factored as **A** = **K S**, where **S** is an $(n$ - 1) × n matrix and **K** is an $\tilde{n}$ × $(n$ -1) matrix. Letting $\mathbf{S} = \left[\mathbf{I}_{n-1} \,\vdots\, -\mathbf{1}_{n-1}\right]$, **K** equals the first n - 1 columns of **A**. Thus, the model can be reparameterized as

$$\boldsymbol{\tau} = -\mathbf{DK}\boldsymbol{\mu}_z \qquad \boldsymbol{P}_{z^*} = \mathbf{D}\left(\mathbf{K}\boldsymbol{\Sigma}_z\mathbf{K}' + \omega^2\mathbf{I}\right)\mathbf{D} \tag{29}$$

where $\mathbf{D} = \left(\text{Diag}\left(\mathbf{K}\boldsymbol{\Sigma}_z\mathbf{K}' + \omega^2\mathbf{I}\right)\right)^{-\frac{1}{2}}$, $\boldsymbol{\mu}_z = \mathbf{S}\boldsymbol{\mu}_t$ is a $(n$ - 1) vector and $\boldsymbol{\Sigma}_z = \mathbf{S}\boldsymbol{P}_t\mathbf{S}'$ is a $(n$ - 1) × $(n$ - 1) matrix. The relation between both sets of parameters is

$$\boldsymbol{\mu}_t^* = \boldsymbol{\mu}_z \qquad \rho_{ii'} = \begin{cases} \dfrac{2 - \tilde{\sigma}_{i'}^2}{2} & \text{if } i = n \\ \dfrac{2 - \tilde{\sigma}_i^2 - \tilde{\sigma}_{i'}^2 + 2\tilde{\sigma}_{ii'}}{2} & \text{otherwise} \end{cases}$$

where we use $\boldsymbol{\mu}_t^*$ to denote the identified parameters in $\boldsymbol{\mu}_t$, and $\rho_{ii'}$ and $\tilde{\sigma}_{ii'}$ to denote the elements of $\boldsymbol{P}_t$ and $\boldsymbol{\Sigma}_z$. Next, we use a Cholesky decomposition for $\boldsymbol{\Sigma}_z$, $\boldsymbol{\Sigma}_z = \mathbf{VV}'$, where **V** is a lower triangular $(n$ - 1) × $(n$ - 1) matrix. Then, to be able to estimate ω (we do not estimate ω^2 directly) we set the $(n$ - 1, n - 1) element of **V** equal to 1. For instance, with $n = 4$ stimuli,

$$\mathbf{V} = \begin{pmatrix} v_{11} & 0 & 0 \\ v_{21} & v_{22} & 0 \\ v_{31} & v_{32} & 1 \end{pmatrix}.$$

The minimization is then performed with respect to $\boldsymbol{\vartheta} = \left(\boldsymbol{\mu}_z{}', \mathbf{v}', \omega\right)'$, where $\mathbf{v} = \left(v_{11}, v_{21} \cdots, v_{n,n-1}\right)'$. These estimates are then transformed to obtain the parameters of interest $\boldsymbol{\theta} = \left(\boldsymbol{\mu}_t^{*\prime}, \boldsymbol{\rho}_t{}', \omega^2\right)'$ and standard errors for $\boldsymbol{\theta}$ are obtained by the multivariate delta method using (22).

A similar approach may be employed to estimate an unrestricted covariance structure model when $\boldsymbol{\Omega}^2$ is assumed to be a diagonal matrix. In this case, the minimization is performed with respect to $\boldsymbol{\vartheta} = \left(\boldsymbol{\mu}_z{}', \mathbf{v}', \boldsymbol{\omega}'\right)'$, where $\boldsymbol{\omega} = \left(\omega_1, \omega_2, \cdots, \omega_{\tilde{n}}\right)'$ and the resulting estimates are transformed to obtain the parameters of interest $\boldsymbol{\theta} = \left(\boldsymbol{\mu}_t^{*\prime}, \boldsymbol{\rho}_t{}', \boldsymbol{\omega}^{2\prime}\right)'$.

NOTAS

Depósito Legal: M-20073 I.S.S.N.: 1579-4873