Psychological Reports, 2007, 100, 211-230. © Psychological Reports 2007

BETWEEN-GROUP DIFFERENCES AND TAXOMETRICS[1]

ISTVÁN HIDEGKUTI AND PAUL DE BOECK

K.U. Leuven, Belgium

Summary.—In psychological research it can be of interest to differentiate between the categorical (taxonic) vs dimensional (nontaxonic) nature of phenomena, for example in the field of psychopathology. For categorical (taxonic) structures a further differentiation is possible: between categorical structures with dimensional equivalence and a substantial main effect, and structures without dimensional equivalence (independent of the main effect). Previous taxometric studies have not concentrated in an explicit way on the case of nonequivalence. It is shown here that under favorable circumstances both types of categorical structures can be detected using the taxometric procedure L-Mode, but that they cannot be differentiated from one another.

In psychology there is a long-lasting debate about the nature of different psychological phenomena. According to some social psychology theories (e.g., Tajfel, 1981) people are inclined to think in terms of categories, thus reducing the information to be processed. Often when categories are used, a continuum is divided into parts or categorical labels are given for the sake of communication. As a consequence, it is sometimes wrongly assumed that a taxonic (i.e., categorical) entity is at the basis of such a "category" (Meehl, 1995). The dimensional vs categorical controversy has stimulated both the systematic investigation of the "true" nature of different phenomena and the development of a methodology.

In psychopathology research, the most popular method to assess whether a phenomenon is dimensional or categorical is *taxometrics*. Taxometrics is a general term covering several procedures. The taxometric methods have been developed by Paul Meehl and his colleagues (e.g., Meehl, 1995; Meehl & Yonce, 1994, 1996; Waller & Meehl, 1998), and over time these methods have gained an increasing popularity and usefulness (see e.g., Haslam & Kim, 2002). The goal of taxometrics is to identify whether the underlying structure of a certain phenomenon is categorical or dimensional. However, the distinction should rather be seen as one between categorical-dimensional and purely dimensional than between categorical and dimensional. "Thus, the convenient taxonomy taxonic-vs-dimensional should, strictly speaking, read 'taxonic-dimensional vs dimensional only' " (Waller & Meehl, 1998, p. 9). Taxometric methods allow one to investigate whether the difference be-

[1]Address correspondence concerning this article to Paul De Boeck, Department of Psychology, K.U. Leuven, Tiensestraat 102, B-3000, Leuven, Belgium or e-mail (paul.deboeck@psy.kuleuven.be). We acknowledge the support from the GOA-00/2 grant from the same university. We are grateful to reviewers for their helpful suggestions.

DOI 10.2466/PR0.100.1.211-230

tween categories is abrupt or smooth. When the differences between groups are smooth, the between-group differences are considered to be dimensional. If they are not, the differences are considered to be categorical.

De Boeck, Wilson, and Acton (2005) have addressed the dimension vs category issue primarily based on a distinction between dimensional equivalence and dimensional nonequivalence, assuming that the structure is often dimensional to some extent as expressed in the quotation from Waller and Meehl (1998, p. 9). Following their DIMCAT framework, there is dimensional equivalence if the categories can be differentiated by the extent of something, e.g., category B having more of the same than category A has. Given dimensional equivalence, a distinction between smooth and abrupt differences is made, corresponding to a small and large difference in extent, respectively. For the structure to be categorical in the case of dimensional equivalence, a large difference (main effect) is required. When the categories cannot be differentiated by the extent of something which is the same in both categories, there is dimensional nonequivalence, in that different dimensions apply to different categories. In other words, in the case of dimensional equivalence the categories are located on the same common latent dimension albeit the difference can be abrupt (steep differentiation, step function). In the case of dimensional nonequivalence the categories do not share a common latent dimension. As mentioned above, in the taxometric framework, the case of dimensional nonequivalence is not described in an explicit way, and it is an interesting issue whether taxometric methods can detect this type of taxonicity as well.

The reason why the dimensional nonequivalence type of taxonicity is not considered in the taxometric research may be that taxometrics is meant for two categories and that ideally the categories are *homogeneous*. Because one can always formally describe the difference between two patterns of values in terms of one dimension (even when the difference between the taxa is qualitative), the difference between taxa (two) can by definition be framed as a matter of extent, albeit an abrupt difference in extent. However, when the categories are heterogeneous, and hence, there are more than two patterns of values, the distinction between dimensional equivalence and dimensional nonequivalence does make sense. Therefore it is important to investigate whether taxometric methods can also detect between-category differences that are of the dimensional nonequivalence type.

In practice, to study a construct, indicator variables are used to assess the construct. When using a dimensional framework (which is a very general framework incorporating also categorical structures) as one tends to do in both DIMCAT and taxometrics, the dimensional location and dimensional relevance of the indicators are the important parameters. In factor analytic terms, these are the mean or threshold and the loading of the indicators,

respectively. Dimensions must be considered the same if the location and the relevance of the indicators are the same (dimensional equivalence). If the locations and/or the relevances differ, the dimensions are different (dimensional nonequivalence). The present study concentrates on location differences, as an initial step and for reasons to be explained later.

To illustrate dimensional equivalence and nonequivalence, two types of hypothetical category pairs will be presented each with four indicators (a, b, c, d) and a positive relation with category membership in comparison with no membership, assuming that the two categories are subsets of a larger set that is not fully covered by the two categories. The first type of pairs (dimensional equivalence) consists of two categories with differences within the categories that can be described with the same dimension for both categories, whereas the two categories differ in their location along the same dimension. Fig. 1(a) illustrates this type. Categories A and B share one underlying dimension, with the same location of the indicators for the two categories (in the order a, b, c, d), but with category B located much higher up along the dimension in comparison with category A, so that they would show as taxa.

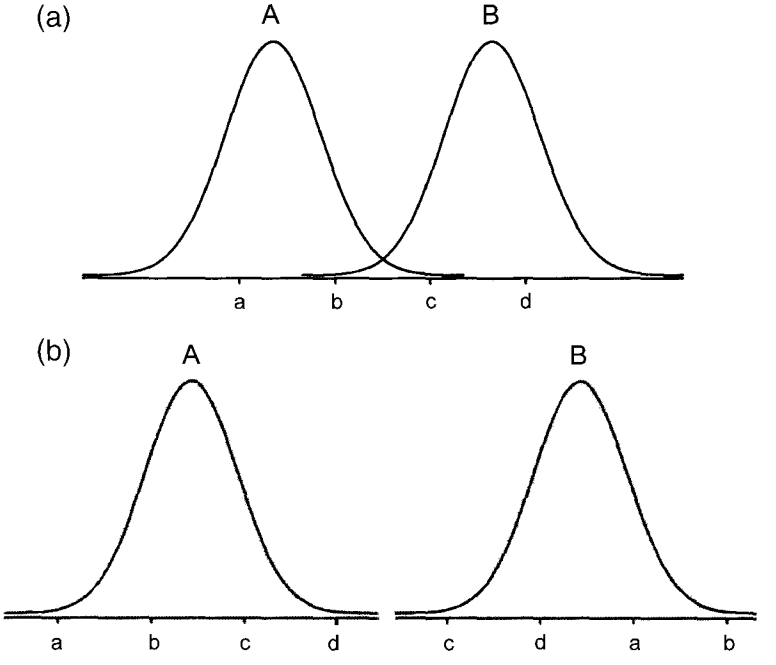


FIG. 1. Graphical representation of dimensional equivalence (a) and nonequivalence (b)

The second type of pairs (dimensional nonequivalence) is illustrated in Fig. 1 (b). It consists of two categories with differences within the categories that cannot be described with the same dimension for both categories, because the indicator locations are different. For example, the order in category A is a, b, c, d, whereas it is c, d, a, b in category B. The difference in location is a difference in mean, or prevalence if symptoms are concerned, for example. In category A, a and b are the most prevalent symptoms, because they have the lowest threshold, whereas in category B, c and d are the most prevalent symptoms.

De Boeck, *et al.* (2005) showed that DIMCAT can detect both kinds of category-likeness (taxonicity): dimensional equivalence with large differences between the means, and dimensional nonequivalence. The issue has not yet been investigated in an explicit way within the taxometric tradition. The aim of the simulation study to be described is to assess the MAXCOV procedure and the L-Mode method from taxometrics on their detection capacity for dimensional nonequivalence, a likely candidate for taxonicity, as shown by De Boeck, *et al.* (2005).

Two Taxometric Procedures

L-Mode: short for Latent Mode (Waller & Meehl, 1998).—This is a relatively new taxometric procedure that is based on factor analysis. Following the L-Mode method, first a single-factor factor analysis is performed and then the factor scores of that single factor are plotted. In case of a taxonic (categorical) construct, the graph of the factor scores is expected to show bimodality, whereas factor scores of nontaxonic (continuous, dimensional) constructs are expected to show unimodality.

Without going into too much detail, L-Mode is based on the notion that factor scores can represent not only continuous underlying variables, but dichotomous ones (such as group membership) as well. [This description is strongly based on Waller & Meehl (1998), where the interested reader can find further details about the procedure.] When the taxonic model holds, that is, when there is a dichotomous underlying variable behind the data, the standardized factor scores (ξ_z) can be expressed as:

$$\xi_z = \frac{s - P}{\sqrt{PQ}} \qquad [1]$$

where s stands for the unstandardized factor score, P is the taxon base rate (the proportion of the sample belonging to the taxon group), $Q = 1 - P$, and $\sqrt{PQ} = \sigma_s$. Given that s denotes a dummy variable with values 0 for nontaxon members and 1 for taxon members, the specific (standardized) factor scores can be written as $\xi_{z_t} = \frac{Q}{\sqrt{PQ}}$ for taxon group members and $\xi_{z_c} = \frac{-P}{\sqrt{PQ}}$ for the complement (nontaxon) group.

Of course, even if the concept underlying the data is taxonic, the factor scores will not consist of only two values because of measurement error, hence, the plot of those scores will form a bimodal distribution. The modes of this bimodal distribution correspond to the factor scores described above and from these the taxon base rate can be expressed as:

$$P = \frac{1}{1 + \xi_{z_t}^2} \qquad [2]$$

for the upper mode (from $\xi_{z_t} = \frac{Q}{\sqrt{PQ}}$), and

$$P = 1 - \frac{1}{1 + \xi_{z_c}^2} \qquad [3]$$

for the lower mode (from $\xi_{z_c} = \frac{-P}{\sqrt{PQ}}$).

MAXCOV: Short for MAXimum COVariance (Meehl & Yonce, 1996).—This technique was the first published taxometric procedure and thus far it is the most popular one among the several methods. It requires two output indicator variables and a third indicator variable as an input indicator which is related to the two categories. The underlying assumption of the MAXCOV procedure is that the output variables are not correlated within a category (i.e., latent homogeneity—no systematic differences exist within the category), but only in a mixed sample of both categories. If this assumption holds,[2] it can be shown that the covariance of the two output indicators along the ordered sections of the input indicator reaches its maximum value in the section where the proportions of taxon and nontaxon group members are equal (at least approximately). The result is a peaked covariance curve. If the input variable is a valid indicator of the phenomenon in question, then the subsamples corresponding to the lowest scores on the input variable will contain (almost) exclusively complement category members, whereas the subsamples corresponding to the highest scores on the input variable will contain (almost) exclusively taxon members (or vice versa). For low values and high values of the input indicator, only a low covariance between the output indicators is expected, whereas for values in the neighborhood of the cut-off between the categories, the covariance is expected to be higher because members of both categories are included and the covariance is assumed to stem from between-category differences if the phenomenon is taxonic.

Because in the hypothetical example of Fig. 1(a) and (b) the sum of the indicators (a + b + c + d) does differentiate between A and B for the first type

[2]Meehl (1995) pointed out that the MAXCOV technique is robust to quite substantial violations of this assumption (a within-group correlation of .30, or, in favorable cases, even higher).

of pairs [Fig. 1(a)], but not necessarily for the second [Fig. 1(b)], it is expected that MAXCOV performs well for the first type, but not for the second unless both types of differences are combined (differences in the average sum of the indicators and in the underlying dimension). The combination of both would require that either for A or for B the distribution is moved upwards or downwards along its category-specific dimension. In other words, if there is no substantial overall group difference (difference in the sum of all indicators) a composite indicator[3] cannot differentiate sufficiently between one category and the other, that is, the composite indicator is not a valid indicator to differentiate between the two categories, whereas it may still be valid as an indicator of membership versus nonmembership in the larger set of persons (e.g., when A and B are two personality disorders in comparison with persons without a disorder). This would result in an approximately equal mix of taxon and complement group members all along the input indicator, so that the covariance curve is expected to be flat.

When sufficiently large, the differences in indicator location between the categories create a negative or very low correlation in the total sample of category A and B between indicators with a clearly different location when going from one category to the other. In the example of Fig. 1(b), a and b will be negatively correlated with c and d or they will be correlated only to a small degree, so that the factor scores for both categories will diverge. Therefore, it is expected that the L-Mode technique can detect dimensional nonequivalence, and hence, differences of the type reflected in Fig. 1(b). A simple sum (a + b + c + d), as would be used for MAXCOV without prior knowledge, would do less well. Of course, if one knew that the sum does not differentiate well between A and B, one would not use the sum as an input indicator. Two remarks apply here. First, the L-Mode method does not require this kind of prior knowledge on the differentiation capacity. Second, it is possible that the sum is a valid indicator for A and B when compared to a control group, for example, two personality disorders in comparison with normality, whereas the comparison under study is the one between A and B as in De Boeck, *et al.* (2005).

Because this is a first study on this issue, the focus is on one type of dimensional nonequivalence only, namely, location differences between indicators, as in Fig. 1(b). This type of difference is not just a theoretical one. In the study reported by Maesschalck (1998) it is shown that the locations of a common set of symptoms are negatively correlated when comparing the borderline personality disorder and the antisocial personality disorder, whereas the relevance (factor loading) is not different in the two categories.

[3]For dichotomous or rating-scale type indicator variables the MAXCOV procedure is mostly carried out using composite input indicators (i.e., the sum of all indicators except the two output indicators under consideration).

The study is focused on the case of questionnaire data: a set of items (indicators) with a rating-scale response format. This is a common type of data in psychology, and also the issue whether a trait or an attitude, as measured in the questionnaire, is dimensional or categorical, is not uncommon. In two additional studies the L-Mode and MAXCOV techniques were used for two different situations: one with only a small number of indicators (smaller than in a questionnaire but more common in taxometric studies) and more response options, and one with largely unequal group sizes.

Method

Simulation Design

Rating-scale data were generated for various types of categories. All data are generated with the Modified Graded Response Model (Muraki, 1990) where the cumulative probability of responding in response category m or higher is defined as:

$$P(Y_{pi} \geq m | \theta_{pc}, c_p) = \frac{\exp(\alpha_{ic}(\theta_{pc} - \beta_{ic} + \omega_m + \gamma_c))}{1 + \exp(\alpha_{ic}(\theta_{pc} - \beta_{ic} + \omega_m + \gamma_c))} \quad [4]$$

and the probability of responding in response category m is

$$P(Y_{pi} = m | \theta_{pc}, c_p) = P(Y_{pi} \geq m | \theta_{pc}, c_p) - P(Y_{pi} \geq m + 1 | \theta_{pc}, c_p) \ , \quad [5]$$

except that for the lowest rating-scale category $P(Y_{pi} \geq m | \theta_{pc}, c_p) = 1$ and for the highest rating-scale category $P(Y_{pi} \geq m + 1 | \theta_{pc}, c_p) = 0$.

- The indices $p\,(p = 1, \ldots, P)$, $i(i = 1, \ldots, I)$, $c(c = 1, \ldots, C)$, and $m(m = 1, \ldots, M)$ refer to persons, indicators, person categories and response categories, respectively.
- Y_{pi} is the response.
- c_p is the category of persons a person belongs to.
- α_{ic} is the discrimination parameter and is set equal to one.
- θ_{pc} is the category-specific latent variable.
- β_{ic} is the category-specific location parameter.
- ω_m is the rating-scale category threshold parameter, separating category m–1 and category m(ω_1 is not used because $P(Y_{pi} \geq m | \theta_{pc}, c_p) = 1$).
- γ_c is the person category (group) difference parameter.

For the case of dimensional equivalence, the subscript c may be omitted.

It has to be noted here that the data are generated based on a specific, although rather general measurement model, whereas Paul Meehl did not assume a specific model in the background of the data to be analyzed with taxometric methods.

The values used to generate the data were as follows:

(a) All data sets contain two groups ($C = 2$) of 200 respondents each.

(b) Eleven indicator variables are used ($I = 11$), each with four points on the rating scale ($M = 4$).

(c) The indicator locations (β_{ic}) range from –2 to 2 by steps of 0.4.

(d) The rating-scale category threshold parameters are –1, 0, and 1, for ω_2, ω_3, and ω_4, respectively.

Three factors were varied in the simulation design, and therefore deemed *design factors*:

(e) The category main effect (γ_c) is 0, 0.5, 1, or 2.

(f) The correlation of the indicator locations is either identical (correlation is 1), independent (correlation is 0) or opposite (correlation is –1) in the two categories. Note that this correlation does not refer to the correlation within the two categories. It is a correlation over indicators but not over persons.

(g) The variance of the normally distributed latent variable (within the categories) is 0, 0.5, or 1. [$\theta \sim N(0, \sigma^2)$, $\sigma^2 = 0$, 0.5, or 1.] Note that as a result, that the average within-category correlations between indicators ranges from .04 to .07 ($\sigma^2 = 0$), from .06 to .09 ($\sigma^2 = 0.5$), and from .12 to .24 ($\sigma^2 = 1$).

Depending on the main effect factor (e), the differences between the categories are either small or large, and depending on the correlation factor (f) dimensional equivalence or dimensional nonequivalence is implemented (correlation of 1, and of 0 or –1, respectively). The category main effect and the correlation of the indicators are considered as the two *difference factors* of the design (they differentiate between the two categories), whereas the variance of the person parameter is considered as a *moderation factor* that is shared by the two categories (although this is not necessary, for the sake of simplicity we assume it is the case). The person locations on the underlying construct (σ_{pc}) are generated either from a single normal distribution (when the group main effect γ_c is 0) or from two normal distributions separated by the group main effect.

The combination of the design factors yields 36 cells. In each of the cells 20 data sets were generated. Of the 36 cells, only those with a main effect of 0 or 2 and a variance of 0 or 1 are discussed here (12 in total). From the results of the other cells, it is clear that they correspond to interpolations between the extreme values of the main effect factor and the variance factor.

Analyses

To perform the L-Mode and MAXCOV analyses, Rusios's (2003) computer software[4] was used. When performing the MAXCOV procedure, sliding windows (50) were employed instead of fixed intervals, because this way

[4]Ruscio, J. (2003) Taxometric programs in R. [Computer software and documentation]. Retrieved September 22, 2003, from http://www.etown.edu/psychology/Faculty/Ruscio.htm.

of proceeding provides more points on the resulting graph, and hence, it provides a more readily interpretable result without making use of smoothing techniques which are considered an ambiguous practice by some researchers (e.g., Ruscio, 2000). Because MAXCOV yields a huge number of graphs (one for each pair of indicators, and hence, 55 in this case) only the summary graph (the average) was used to evaluate the MAXCOV results. However, we will also briefly describe the results obtained for the pairs.

Results

Fig. 2 depicts the typical L-Mode plots for the 12 cells. Each row in Fig. 2 contains three plots, and each plot refers to one data set and gives a representative view on the results.[5] First, the focus is on the homogeneous categories (variance is 0, first two rows), then the moderation effect of within-category heterogeneity (variance is 1, last two rows) is discussed.

In row 1, L-Mode plots of the data sets with homogeneous categories and no main effect (main effect is 0) are shown. The only factor varied is the amount of correlation between the indicator locations. The case of dimensional equivalence is shown in plot a. As can be seen in plot a, the L-Mode analysis provides a unimodal curve indicating that there are no differences between the categories.

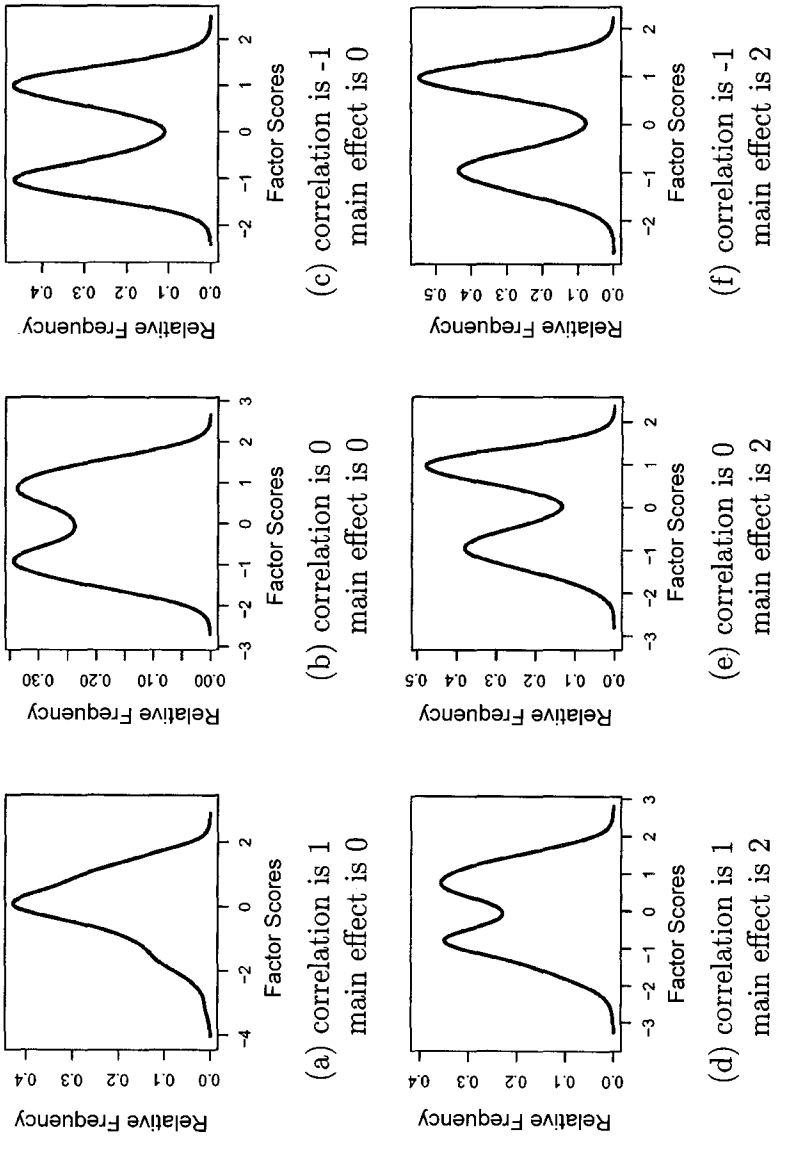
When there are location differences involved (i.e., dimensional nonequivalence), as is the case in plots b and c in row 1, the L-Mode shows bimodality although the category main effect is zero. Note that the bimodality is more apparent for a negative correlation than for a zero correlation (going from left to right). When the main effect is 2 as in row 2, the L-Mode again correctly indicates taxonicity. It does so to an extent that depends on the location differences, but also for identical indicator locations the L-Mode shows the two categories, although to a much smaller degree.

In sum, for homogeneous categories the L-Mode technique does detect both types of between-group differences: main-effect differences as well as dimensional nonequivalence. One of both in the absence of the other is sufficient for the L-Mode to detect taxa, but a combination of both yields a more extreme differentiation.

It is no surprise that the results obtained with heterogeneous categories (rows 3 and 4) are somewhat different from those obtained with homogeneous categories. As expected,[6] heterogeneity moderates the effects of the two difference factors, but the moderation is stronger when the main effect is concerned than when nonequivalence is concerned. The curves show bi-

[5]All the plots available upon request from the first author.

[6]Within-category heterogeneity means that the indicators do covary in the categories, violating the assumption of zero nuisance covariance, hence creating less favorable circumstances for the procedure.


Relative Frequency
Factor Scores
(a) correlation is 1
main effect is 0
(b) correlation is 0
main effect is 0
(c) correlation is -1
main effect is 0
(d) correlation is 1
main effect is 2
(e) correlation is 0
main effect is 2
(f) correlation is -1
main effect is 2

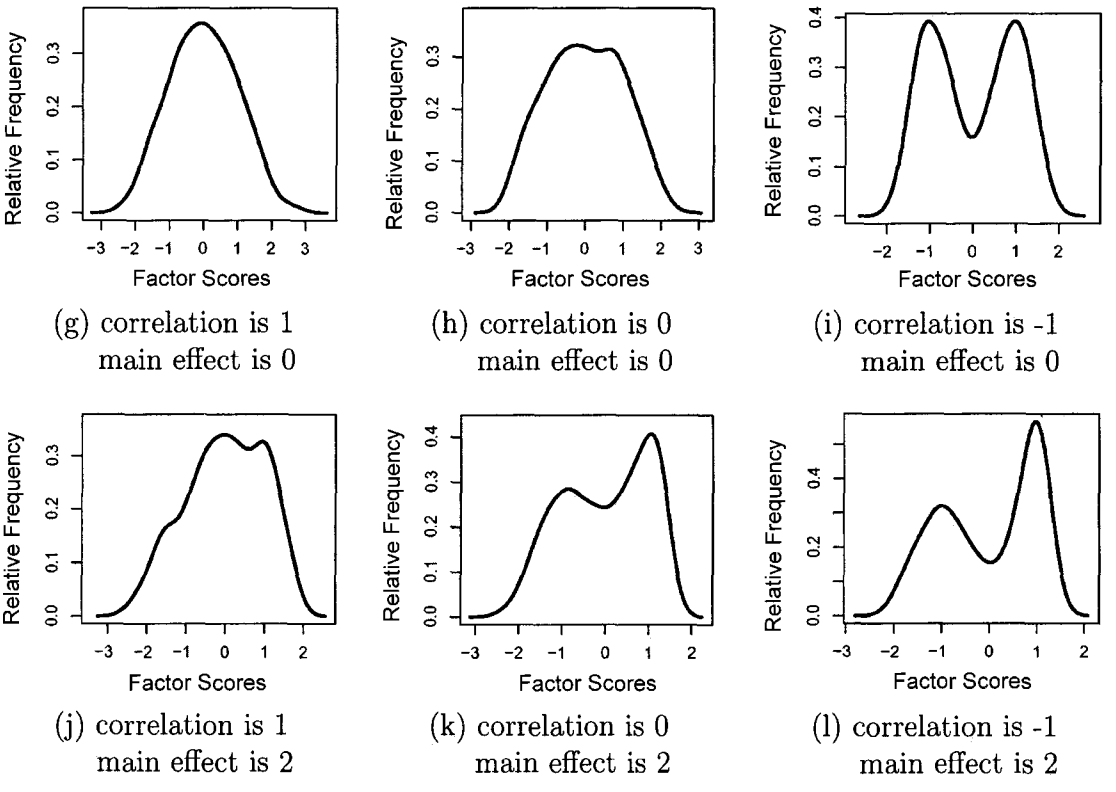
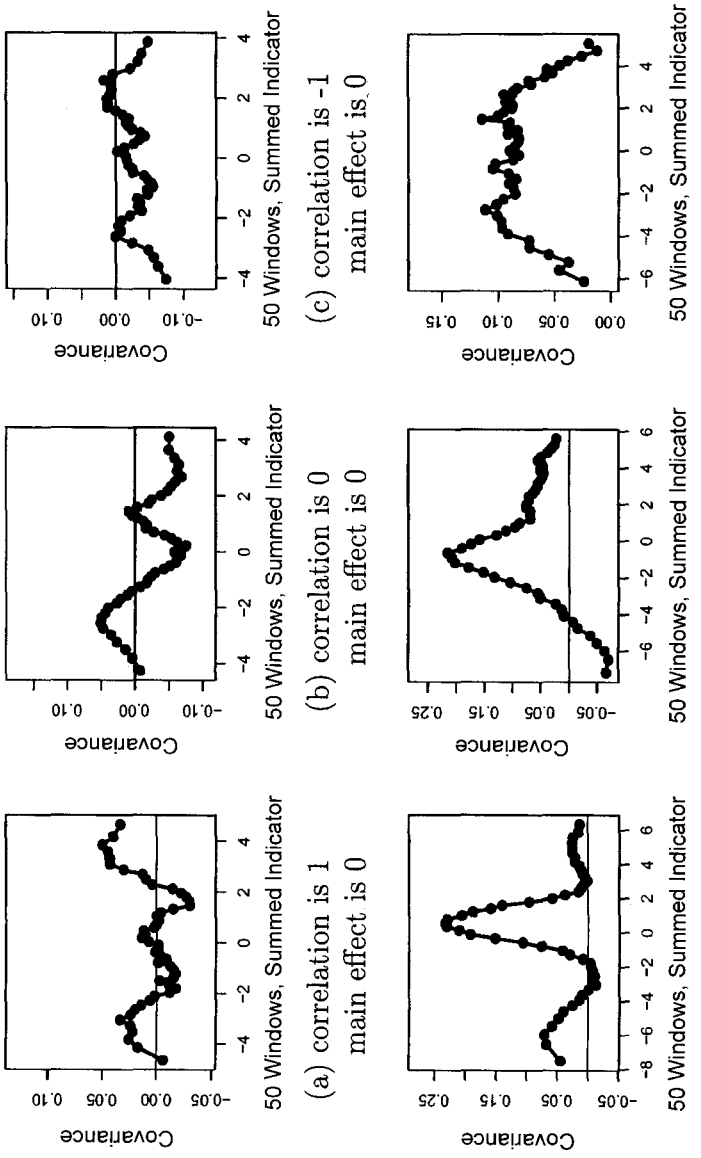


FIG. 2. L-Mode graphs for homogeneous categories (a to f) and heterogeneous categories (g to l) depending on dimensional equivalence and main effect.


Covariance
50 Windows, Summed Indicator
(a) correlation is 1
main effect is 0
(b) correlation is 0
main effect is 0
(c) correlation is -1
main effect is 0
(d) correlation is 1
main effect is 2
(e) correlation is 0
main effect is 2
(f) correlation is -1
main effect is 2

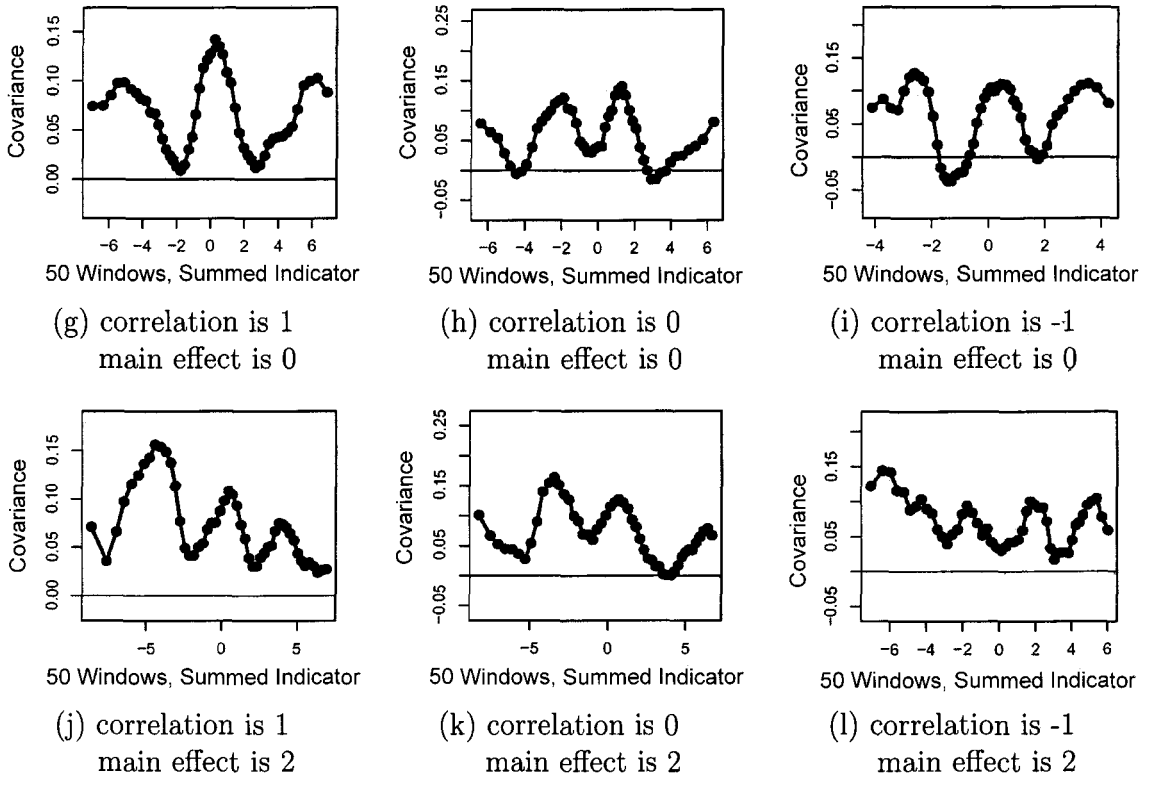


FIG. 3. MAXCOV graphs for homogeneous categories (a to f) and heterogeneous categories (g to l) depending on dimensional equivalence and main effect.

modality clearly only when the locations are negatively correlated, and a main effect of 2 is not sufficient to detect that there are two categories. Clearly, the L-Mode technique is sensitive to within-category heterogeneity along the underlying dimension. An even larger category separation than provided by a main effect of 2 or an outspoken dimensional nonequivalence (i.e., correlation of –1) are required for categories to be detected when they are heterogeneous. The detection of nonequivalence seems much easier, which shows that the L-Mode approach is quite appropriate to detect nonequivalence as the basis of a categorical structure.

In Fig. 3 the corresponding averaged MAXCOV graphs are presented. The peaked curve which is typical for taxonicity shows clearly only for homogeneous categories and when the main effect is 2, while the correlation between the locations is positive or zero, as in plots d and e in row 2. In the other cells, the curves show very low covariance or have multiple peaks. Based on the inspection of the individual (not averaged) MAXCOV graphs, one would come to the same conclusion as with the inspection of the averaged graphs. Clearly, when the sum of the indicators is used as an input indicator, the MAXCOV technique can detect taxonicity only when the main effect is rather large and the categories are homogeneous while the locations are not negatively correlated.

Two Additional Studies

In the main simulation study the focus was on a questionnaire-like arrangement of the data, that is, a relatively large number of indicators (11) were employed with a relatively small number of response options of a rating scale. We believe that this arrangement is a relevant one, however, there are concerns which are to be addressed in two additional studies.

First additional study.—The first concern with the data setup is that it does not follow the classical taxometric data setting, where usually only a few indicators are used (four in most simulation studies). For this reason a simulation study was conducted in which the original study was repeated except for two features, the number of indicators was reduced from eleven to four ($I=4$) and the number of response options was changed from four to eight ($M=8$), so that the data arrangement is more representative of actual taxometric studies. Also the way of analysis was altered in that the analyses were completed using indicator triplets (one input, two outputs) instead of summed indicators. (The increased number of response options makes it reasonable to use single indicators as scales.) The results of the first additional study appear in Figs. 4 and 5 (representative examples for L-Mode and MAXCOV, respectively). To keep matters simple only homogeneous groups were considered in both additional studies.

Based on the outcome of this study, the basic findings are the same as

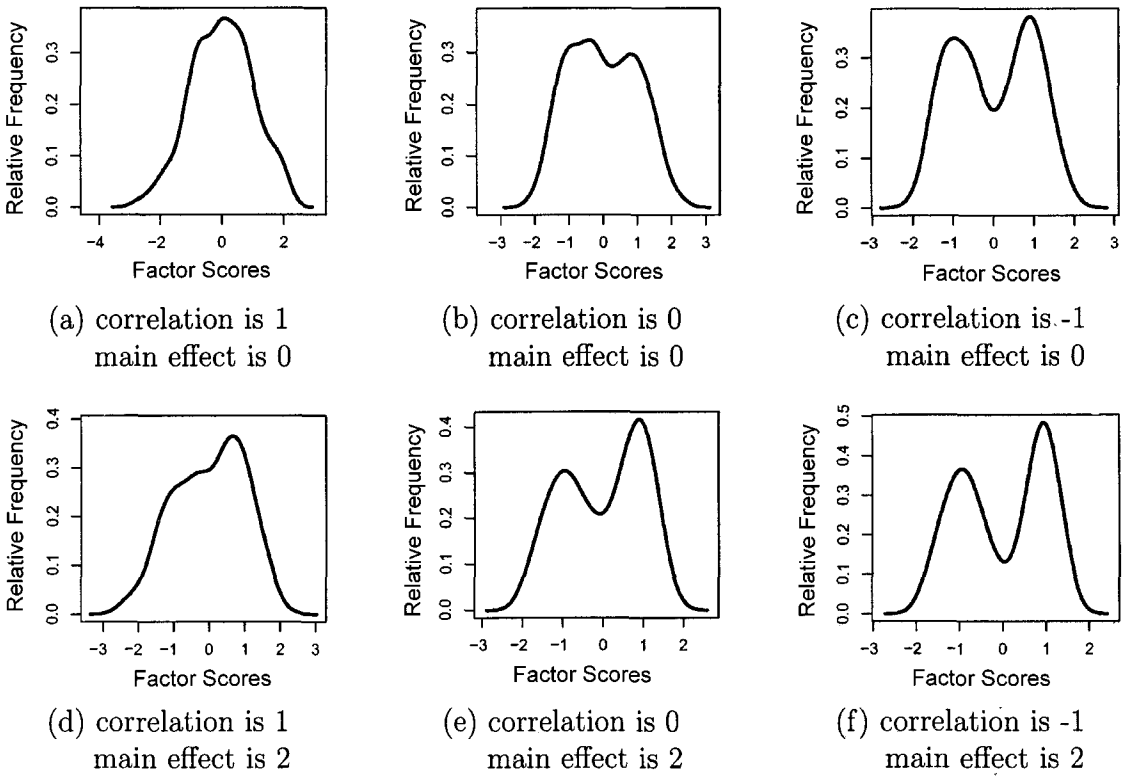


FIG. 4. L-Mode graphs for homogeneous categories depending on dimensional equivalence and main effect. (First additional study)

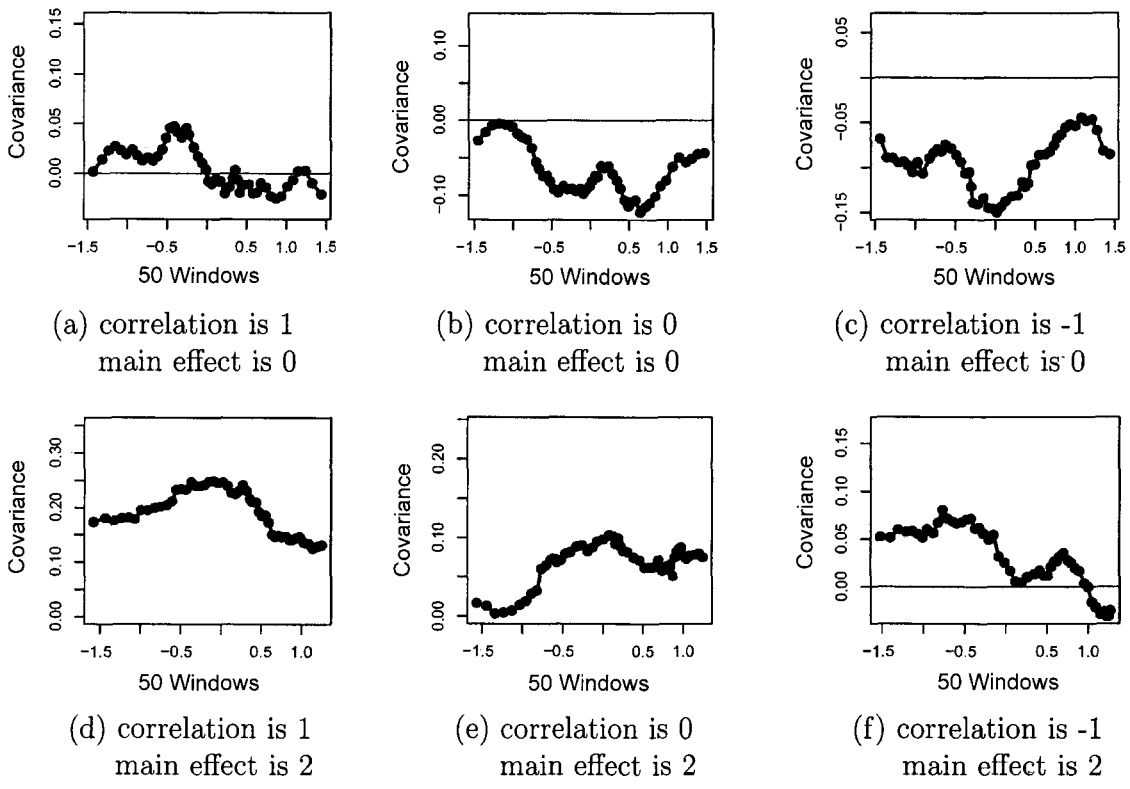


FIG. 5. MAXCOV graphs for homogeneous categories depending on dimensional equivalence and main effect. (First additional study)

in the original study, suggesting that the results do not depend too much on the data arrangement. The results seem to generalize to data that are reasonably representative of actual taxometric studies.

Second additional study.—The second concern is that in the original study the proportions of the groups are equal, that is, the base rate is .50, which is rarely the case in real data settings, where small base rates for target groups are common. To investigate whether smaller base rates of target groups do affect the conclusions, the original study was repeated but now with unequal group sizes. Data were generated for 100 respondents in one group and 300 respondents in the other group, corresponding to a base rate of .25 for the first group. As can be seen from Figs. 6 and 7 (representative examples for L-Mode and MAXCOV, respectively) the basic conclusions of the original study generalize also to a smaller base rate.

Discussion

The results of the simulation studies show that the L-Mode technique can detect both types of category-likeness (taxonicity), a large main effect and dimensional nonequivalence, and that the heterogeneity of the categories has a negative but not detrimental effect on the clarity of the detection of dimensional nonequivalence between the categories. On the other hand, the MAXCOV technique can detect taxa in this study only in certain conditions: a sufficiently large main effect, homogeneous categories, and not too strong deviations from dimensional equivalence.

In sum, the two types of category-likeness, a large main effect and dimensional nonequivalence, can be detected by the L-Mode technique. This means that a type of taxon which was not earlier investigated in an explicit way within the taxometric approach can be dealt with by taxometrics as well. In principle, the application of the MAXCOV technique can be adapted in order for the method to detect not only a large main effect but also dimensional nonequivalence as well. The condition is that one has prior knowledge of the prevalence order of the indicators in the groups under consideration, so that a weighted sum of the indicators can be used (as in a factor score such as for the L-Mode technique), whereas for the L-Mode technique no such prior knowledge is required. There are two cases where such knowledge would not be available: when the features of the groups are not known in advance or when it is even not known which groups are involved. The prior knowledge necessary for MAXCOV to detect dimensional nonequivalence may stem from an application of the L-Mode method.

The simulation study has several limitations. First, a specific model is used to generate the data. Albeit the model used in the present study is a rather common model, it may be of interest to use also other models. Second, dimensional nonequivalence is restricted to differences in location and

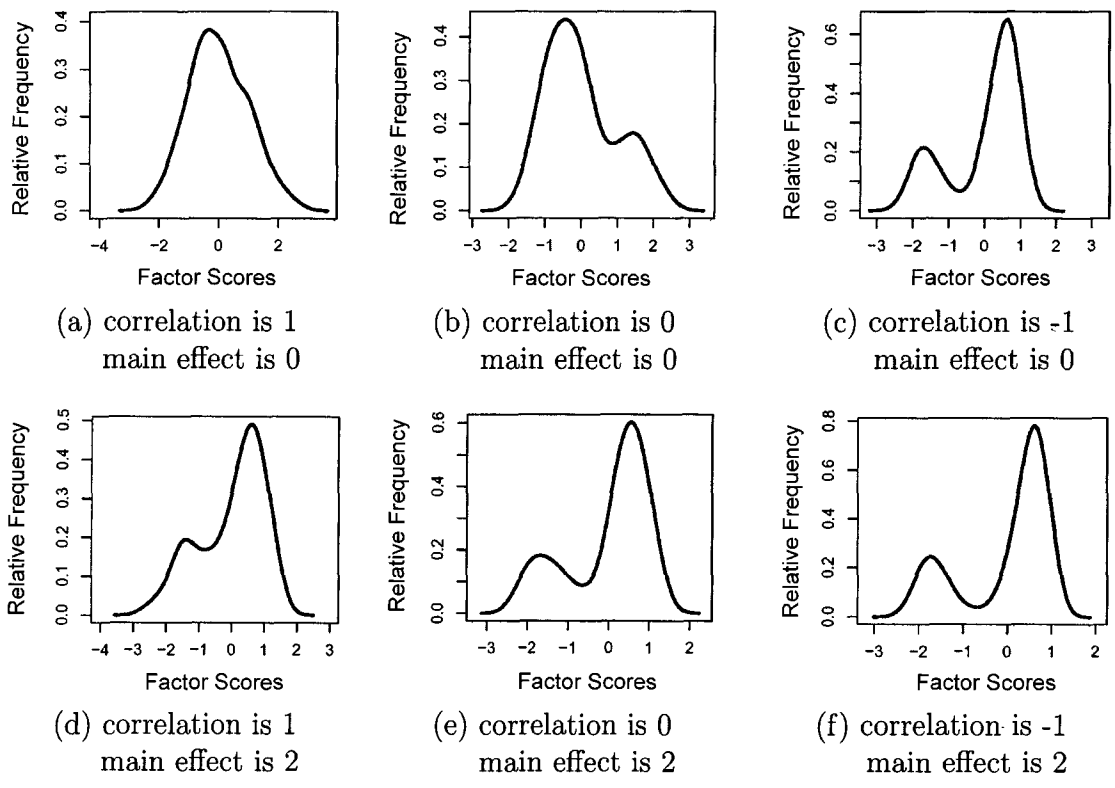


FIG. 6. L-Mode graphs for homogeneous categories depending on dimensional equivalence and main effect. (Second additional study)

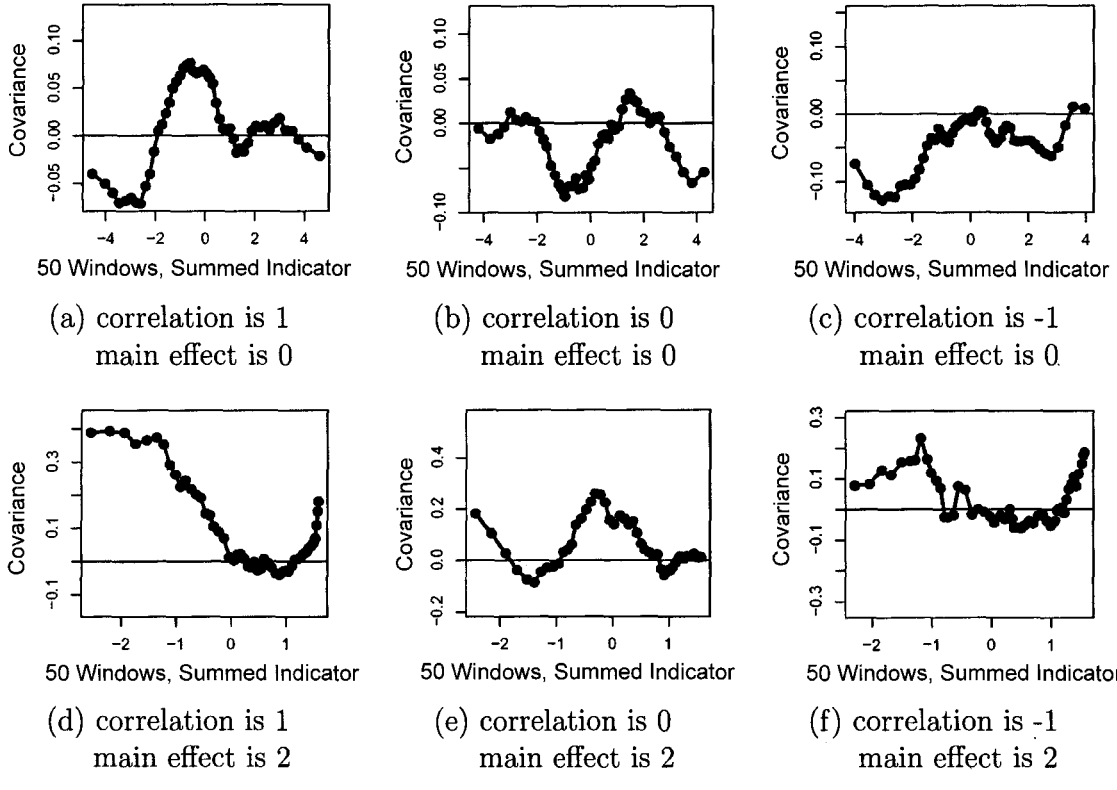


FIG. 7. MAXCOV graphs for homogeneous categories depending on dimensional equivalence and main effect. (Second additional study)

thus in the order of the means between the categories. The results may be different (and presumably more complicated) for differences in relevance and thus in factor loadings within the categories. For example, when the difference between the categories is between positive relevance and no relevance, the resulting structure is one with a homogeneous category and a heterogeneous category, since lack of within-category relevance implies homogeneity at the latent level and true heterogeneity implies within-category relevance. This complication is an additional reason why it made sense to concentrate first on location nonequivalence as in the present study. Third, only a limited set of data configurations is used, and these were also somewhat idealized, such as the assumption of a normal distribution of the latent trait within the categories. Other, less ideal data configurations should also be a subject of further research. The values of the design factors are not chosen to be gradual. For example, only three values of heterogeneity were chosen (0, 0.5, 1), whereas it would be of interest to investigate also more values, so that one may assess in a more fine-grained way how heterogeneous the categories can be for MAXCOV to still detect taxonicity. Fourth, the data are rating-scale data. It may be of interest to investigate also the capacity of the taxometric methods to detect taxa of the dimensional nonequivalence type using quasi-continuous and possibly also dichotomous indicators. The use of dichotomous variables in taxometric research has been criticized (e.g., Miller, 1996) and needs to be further explored.

REFERENCES

De Boeck, P., Wilson, M., & Acton, S. (2005) A conceptual and psychometric framework for distinguishing categories and dimensions. *Psychological Review*, 112, 129-158.

Haslam, N., & Kim, H. C. (2002) Categories and continua: a review of taxometric research. *Genetic, Social and General Psychology Monographs*, 128, 271-320.

Maesschalck, C. (1998) A psychometric modeling framework for testing categorical and/or dimensional aspects of the borderline, histrionic and antisocial personality disorders. Unpublished doctoral dissertation. K. U. Leuven, Belgium.

Meehl, P. E. (1995) Bootstraps taxometrics, solving the classification problem in psychopathology. *American Psychologist*, 50, 266-275.

Meehl, P. E., & Yonce, L. J. (1994) Taxometric analysis: I. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAXCOV procedure). *Psychological Reports*, 74, 1059-1274.

Meehl, P. E., & Yonce, L. J. (1996) Taxometric analysis: II. Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator (MAXCOV procedure). *Psychological Reports*, 78, 1091-1227.

Miller, M. B. (1996) Limitations of Meehl's MAXCOV-HITMAX procedure. *American Psychologist*, 51, 554-556.

Muraki, E. (1990) Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, 14, 59-71.

Ruscio, J. (2000) Taxonomic analysis with dichotomous indicators: the modified MAXCOV procedure and a case removal consistency test. *Psychological Reports*, 87, 929-939.

Tajfel, H. (1981) *Human groups and social categories: studies in social psychology*. Cambridge, MA: Harvard Univer. Press.

Waller, N. G., & Meehl, P. E. (1998) *Multivariate taxometric procedures: distinguishing types from continua*. London: Sage.

Accepted January 19, 2007.