Detecting Response Styles by Using Dual Scaling of Successive Categories

Michel van de Velden *

October 12, 2007

Econometric Institute Report EI 2007-41

Abstract

A response style denotes a certain mapping of latent preferences to a rating scale that is common among a certain group of individuals. For example, individuals from the same country may assign high ratings to the majority of objects regardless of the specific preferences for the objects. The existence of response styles causes problems in international and cross-cultural research as it makes it hard to compare findings. Moreover, even within homogeneous samples, response styles make it difficult to expose the underlying preference structure. Detecting the existence and influence of a response style is typically a difficult issue as the underlying preferences are not directly observable. Hence, we can never be sure if the observed ratings are the result of a response style or an adequate representation of the preferences. In this paper, we consider the use of dual scaling as a tool to detect the existence of a response style. By means of a simulation study, we assess the performance of the proposed method.

1 Introduction

For the analysis of attitudes or preference, researchers frequently use rating data. In rating data, objects or attributes are evaluated by individuals on a, usually predefined, scale. For example, in marketing research, a researcher may ask individuals to rate a set of soft drinks or attributes of these softdrinks according to their preferences. Rating data are a popular format for several reasons: Subjects generally understand the task without great effort, the ratings are easy to administer and there exist many methods to analyse the ratings. Popular methods for the analysis of rating data are principal component analysis, unfolding analysis, dual scaling or correspondence analysis.

*Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (e-mail: vandevelden@few.eur.nl)

When analyzing rating data, one often implicitly assumes that the ratings can be regarded as interval data. That is, differences between ratings have equal meaning across individuals. This, however, may not always hold. In particular, in studies where individuals come from different backgrounds, e.g. different nationalities, cultures or levels of education, ratings may have different meanings across individuals. For example, individuals with a certain background may be inclined to give high ratings regardless of their true underlying preferences. Such behavior is usually referred to as a response style. The existence of response styles may severely influence the results of an analysis that does not acknowledge their existence.

A response style can be seen as a, possibly non-linear, mapping of the underlying latent preferences to a rating scale that is common among a group of individuals. For example, for a group of individuals, there may be a tendency to give high ratings to all items regardless of the actual latent preferences. Consequently, the obtained ratings do not immediately reflect the latent preferences that one wishes to measure. The existence of response styles has been acknowledged in several areas (e.g., Marin, Gamba and Marin, 1992, Baumgarter and Steenkamp, 2001, Van Herk, Poortinga and Verhallen, 2004). However, as the latent preferences are unobservable, one can never disentangle the true preferences from the observed ratings. It is therefore difficult to detect a response style.

In this paper, we consider dual scaling of rating data to detect response styles. In particular, we consider dual scaling of successive categories. Dual scaling of successive categories is a mathematically simple method that yields optimal scaling values for items as well as for inter-rating boundaries. By observing the scores for the inter-rating boundaries, we may be able to detect certain response styles. To assess the performance of the method, we use a simulation study. In the simulation study we are able to generate ratings that are the result of certain response styles. Moreover, by generating data, we are able to explicitly consider "true" preferences as well as the observed ratings. Hence, we are able to see whether, and under which conditions, dual scaling can be used to detect a response style.

The remainder of this paper is organized as follows. In the next section, the dual scaling of successive categories methodology is briefly explained. The simulation study and its results are described in Section 3 and we finish with concluding remarks and suggestions for future research.

2 Dual Scaling of Successive Categories

Dual scaling (Nishisato, 1980) is a versatile multivariate method that is typically used for the analysis of categorical data. The method is closely related and mathematically equivalent to correspondence analysis (Greenacre, 1984). However, for the analysis of preference data the two methods are not the same. First of all, Nishisato (1994) explicitly distinguishes between three types of preference data; paired comparison data and rank order data, rating data. In correspondence analysis on the other hand, only rating data are treated explicitly. Secondly, for the three types of data, Nishisato proposes to transform the raw data to a so-called dominance matrix. Dual scaling is then applied to the dominance data. As the dominance matrix contains negative entries, an impossibility in ordinary dual scaling, the usual method must be applied in a modified fashion. See Nishisato (1994) for details on this modified approach. In correspondence analysis a completely different route is taken: The original rating data are doubled with respect to the items and an ordinary correspondence analysis is applied to the doubled matrix (e.g., Benzécri, 1973, Greenacre, 1984). Although the "doubling" approach appears to differ considerably from the approach proposed by Nishisato, Van de Velden (2000), and Torres and Greenacre (2002), showed that Nishisato's method can also be expressed as dual scaling of a doubled data matrix. There are, however, important differences between the two approaches, see van de Velden (2003) for a complete treatment of these differences.

In addition to the mathematical differences, which have been clarified in the aforementioned references, Nishisato's treatment of rating data differs from the correspondence analysis approach with respect to the specific coding that he proposes. In correspondence analysis it is customary to use the ratings as they are. Hence, they are implicitly assumed to be interval data. Nishisato, on the other hand, proposes two strategies: 1) Compare the ratings for the objects and recode the data as paired comparison data (i.e., count how many times objects received higher ratings than other objects). 2) Recode the data by using successive categories before applying dual scaling. In this paper, we shall concern ourselves with the latter approach.

2.1 Successive Categories

Dual scaling of successive categories is based on the work by Nishisato (1980) and Nishisato and Sheu (1984). The successive categories approach is best explained by means of a small example. Suppose that three objects are to be rated on a 1-5 scale. The scale can then be represented as a line with boundaries between the scale values. Hence, for this example, we have 4 boundaries. Now, suppose that the three objects, say A, B, and C, received ratings 1, 2, and 5 respectively. We can then place the objects on the scale. This is depicted in Fig. 1, where b1, ..., b4 denote the boundaries. The data can be collected in a matrix where the first 3 columns correspond to the objects, the next 4 columns correspond to the boundaries and the rows correspond to individuals. The ijth cell element corresponds to the number of times that, for individual i, an object or boundary is greater than the other objects and boundaries. In other words, boundaries and objects are rank ordered. When equal ratings are assigned to objects, the tied objects are assigned the average rank number. For the example data, we obtain the row: [0 2 6 1 3 4 5]. It is not difficult to generalize the example to the case with p objects and a rating scale ranging from 1 to q. Nishisato (1994, p.223) gives an algorithm to transform the ratings to successive categories.

Dual scaling of successive categories yields optimal scaling values for the

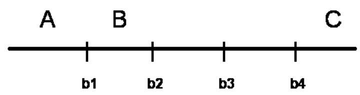

Figure 1: Successive categories

objects and the boundaries. If there exists a tendency to mainly use, for example, the end-points of the scales, this leads to a relatively large gap between the end-point boundaries and small gaps between the middle rating boundaries. Similarly, if the ratings are mapped predominantly to the middle of the scale, the higher boundaries are closer to each other whereas there is a bigger gap between the middle boundaries. This property of dual scaling may be exploited to detect response styles.

3 Simulation Study

To see whether the differences between boundaries can be used to expose a response style, we use a simulation study. In the simulation study, we randomly generate preferences and map these preferences to a rating scale. For the mapping of the generated ratings, we use different functions representing different response styles. Than, by applying dual scaling of successive categories to the data, we obtain dual scaling scores for objects and boundaries. If no response pattern exists, the distances between the boundaries should be approximately equal. Therefore, deviations from this equal spaced pattern may indicate the existence of a response style. To assess the deviation from the equal spaced pattern, we first calculate the difference between the scores for the highest and lowest boundaries and divide this by the number of boundaries minus one (that is, the number of intervals between the boundaries). This is the equal space distance between boundaries. Note that this equal space distance differs from data set to data set. To account for this, we shall consider the deviation with respect to the equal space differences. We call this measure the *relative inter-boundary deviations*. To calculate the relative inter-boundary deviations we divide the differences between the obtained inter-boundary differences and the equal space distances through the equal spaced distances. If the boundaries are equally spaced, the relative inter-boundary deviations are equal to zero.

3.1 Data Generating Process

We consider a situation in which 5 objects are to be ranked on a 1-7 scale. First we draw the underlying preference structure from a uniform [0,1] distribution. We denote this latent structure by μ. We generate preferences for individuals by adding noise to the underlying latent preference μ. The noise is generated by drawing from a normal distribution and multiplying this by a factor α. In the simulations, we vary the α parameter to study its influence on the outcomes. The thus obtained preferences are mapped to a 1-7 rating scale. For this map-

Table 1: Rating Functions; The rows give the assigned ratings, the columns indicate the corresponding intervals of the generated preferences

	No response Style	Aquiescence	Midpoint	Extreme
Rating 1	$[-\infty,1/7)$	$[-\infty,0.10)$	$[-\infty,0.02)$	$[-\infty,0.35)$
Rating 2	$[1/7,2/7)$	$[0.10,0.15)$	$[0.02,0.05)$	$[0.35,0.45)$
Rating 3	$[2/7,3/7)$	$[0.15,0.18)$	$[0.05,0.20)$	$[0.45,0.49)$
Rating 4	$[3/7,4/7)$	$[0.18,0.20)$	$[0.20,0.80)$	$[0.49,0.51)$
Rating 5	$[4/7,5/7)$	$[0.20,0.40)$	$[0.80,0.95)$	$[0.51,0.55)$
Rating 6	$[5/7,6/7)$	$[0.40,0.70)$	$[0.95,0.98)$	$[0.55,0.65)$
Rating 7	$[6/7,\infty)$	$[0.70,\infty)$	$[0.98,\infty)$	$[0.65,\infty)$

ping of the "true" preferences to the rating scale we consider four functions. The first function corresponds to the situation in which no response style exists. We model this by dividing the [0,1] interval into 7 equal sized intervals. Ratings are then assigned by considering the interval in which the true preference falls. Note that the true preference may be negative or larger than 1 and we assign the ratings 1 and 7 to such preferences respectively. The second function corresponds to the so-called *aquiescence* response style. In this response style there is a tendency towards the positive ratings. The third function represents an *extreme* response style in which there is a tendency to assign either high or low ratings. Finally, the fourth function we consider represents the *midpoint* response style. In this response style there is a tendency to assign the middle rating. For more details on these, and other, response styles see, for example, Greenleaf (1992). The exact functions that we use to model the response styles as well as the no response style mapping, are collected in Table 1 The rows of Table 1 give the assigned ratings and the elements indicate to which preference values these ratings correspond for the three response styles.

3.2 Design of the experiment

We consider two experiments to study the ability of dual scaling to detect a response style. In the first experiment, we randomly generate 50 latent preference structures μ and for each randomly generated preference structure μ and noise level α, we generate 20 data matrices. We then transform the "true" preferences to ratings by using the functions summarized in Table 1. To obtain the scores for the boundaries, we apply dual scaling of successive categories to the ratings. Finally, using the dual scaling results, we calculate the relative inter-boundary deviations.

The second experiment is concerned with the question whether it is possible to distinguish an actual response style from a true preference structure that resembles a response style. To study this question we conduct a simulation where we fix the preferences to the three response styles used in this paper. That is, preference structure 1, with as mean vector $\mu_a = [0\ 1\ 1\ 1\ 1]$ corresponds

Table 2: Relative Inter-Boundary Deviations: No Response Style

	α=1/4		α=1/3		α=1/2	
Boundaries	Mean	Std	Mean	Std	Mean	Std
1-2	-0.03	0.13	-0.03	0.09	-0.02	0.05
2-3	0.02	0.08	0.02	0.05	0.01	0.03
3-4	0.03	0.04	0.03	0.02	0.02	0.01
4-5	0.02	0.08	0.02	0.05	0.01	0.03
5-6	-0.04	0.12	-0.04	0.08	-0.03	0.05

Table 3: Relative Inter-Boundary Deviations: Aquiescence Response Style

	α=1/4		α=1/3		α=1/2	
Boundaries	Mean	Std	Mean	Std	Mean	Std
1-2	-0.24	0.05	-0.23	0.04	-0.19	0.02
2-3	-0.28	0.04	-0.26	0.03	-0.22	0.01
3-4	-0.30	0.03	-0.28	0.02	-0.24	0.01
5-4	0.23	0.10	0.22	0.07	0.19	0.03
5-6	0.59	0.16	0.55	0.10	0.46	0.05

to the Aquiescence style. Similarly, let $\mu_e = [1\ 1\ 0.5\ 0\ 0]$ and $\mu_m = [0.5\ 0.5$ $.5\ 0.5\ 0.5]$ denote preference structures for the extreme and midpoint styles respectively. Now, using these vectors as underlying preference structure, we simulate, for each structure and noise level α, 100 data sets. The thus obtained true preferences are mapped to the rating scale without using a rating style. That is, we use the first column of Table 1 to map the preferences to ratings. We then apply dual scaling of successive categories.

3.3 Results

For each μ we have 20 observation matrices. We first calculate the mean relative inter-boundary deviations over these 20 observations. We then calculate the mean and standard deviation of these averages over the 50 simulations. The thus obtained means and standard deviations are collected in Tables 2 through 5. Each Table corresponds to the dual scaling results for a certain response style.

The results in Tables 2 through 5 indicate that the response styles may indeed be detected by dual scaling of successive categories. For example, for the aquiescence style (Table 3) we see that the inter-boundary differences between the 5th and 6th boundary (i.e. between ratings 6 and 7) is much larger than the differences between the 1st and 2nd boundaries. Recall that a value of zero would indicate no deviation from the equal spaced pattern, whereas a value of 0.5 indicates that the deviation was half the equal spaced difference. In

Table 4: Relative Inter-Boundary Deviations: Extreme Response Style

	α=1/4		α=1/3		α=1/2	
Boundaries	Mean	Std	Mean	Std	Mean	Std
1-2	0.14	0.06	0.13	0.04	0.11	0.03
2-3	-0.07	0.01	-0.06	0.01	-0.06	0.01
3-4	-0.15	0.02	-0.14	0.01	-0.11	0.01
4-5	-0.07	0.01	-0.06	0.01	-0.05	0.01
5-6	0.14	0.06	0.13	0.05	0.11	0.03

Table 5: Relative Inter-Boundary Deviations: Midpoint Response Style

	α=1/4		α=1/3		α=1/2	
Boundaries	Mean	Std	Mean	Std	Mean	Std
1-2	-0.40	0.04	-0.37	0.03	-0.33	0.02
2-3	-0.15	0.13	-0.14	0.09	-0.11	0.05
3-4	1.09	0.16	1.03	0.10	0.89	0.05
4-5	-0.15	0.12	-0.14	0.09	-0.11	0.05
5-6	-0.40	0.04	-0.37	0.03	-0.33	0.02

other words, the inter-boundary difference for the highest boundaries was 1.5 times the equal space difference. For the extreme response style, we see that the differences between the 1st and 2nd as well as between the 5th and 6th boundaries are larger than the equal space difference. The difference between the other boundaries, in particular between 3 and 4, are smaller than the equal space difference. For the midpoint response style, a large gap surfaces between boundaries 3 and 4 due to the high number of individuals assigning the middle rating. Finally, as shown in Table 2, when no response style is employed, the relative deviations are close to zero.

For all response styles we see that the surfacing of the response patterns becomes stronger if less noise is added. That is, for smaller values of α the absolute values of the relative deviations increase. At the same time, the standard deviations of the relative deviations increase indicating higher variability in the boundaries when less noise is added. This is porbably due to the fact that the preference structures are randomly drawn from a uniform (0,1) distribution. For each structure we draw 20 samples. By decreasing the noise, these samples become more homogeneous. Consequently, the original variability stemming from the uniform distribution, plays a larger role when α increases.

The results for the second experiment are reported in Tables 6 through 8. We see that, especially when the noise level is relatively high; $\alpha = 1/2$, the inter-boundary differences show no clear response styles for all three preference structures . For the extreme response structure, this situation is hardly influenced by the reduction of noise. For the other two patterns, we see that if less

Table 6: Relative Inter-Boundary Deviations: Aquiescence Preferences

	α=1/4		α=1/3		α=1/2	
Boundaries	Mean	Std	Mean	Std	Mean	Std
1-2	-0.09	0.03	-0.11	0.03	-0.08	0.03
2-3	-0.13	0.03	-0.10	0.03	-0.05	0.03
3-4	-0.10	0.03	-0.04	0.04	-0.01	0.04
4-5	0.03	0.04	0.05	0.04	0.05	0.04
5-6	0.28	0.05	0.19	0.05	0.09	0.05

Table 7: Relative Inter-Boundary Deviations: Extreme Preferences

	α=1/4		α=1/3		α=1/2	
Boundaries	Mean	Std	Mean	Std	Mean	Std
1-2	0.03	0.04	0.01	0.04	-0.01	0.04
2-3	-0.01	0.04	-0.00	0.04	0.01	0.04
3-4	-0.04	0.04	-0.02	0.04	0.00	0.04
4-5	-0.01	0.04	-0.01	0.04	0.01	0.04
5-6	0.03	0.04	0.01	0.04	-0.01	0.04

noise is added, i.e., if the data are closer to the mean patterns μ_a or μ_m, the response styles do surface, albeit less pronounced as previously.

Note that although the dual scaling solution exhibits more or less equal spaced boundaries, the objects receive scores that are in accordance with the underlying preferences. Hence, for the aquiescence response style, the first four objects receive similar, high, scores whereas the other object receives a distinctively lower score. These results are not reported here for the sake of brevity.

The role played by the noise level is crucial for detecting response styles. When no response style exists, the noise is distributed more or less evenly across the ratings. Consequently, the boundaries will be spread out. On the other hand, when noise is mapped according to a response style, the spacing between the boundaries is influenced according to the response style and will surface as such.

4 Conclusion

In this paper, dual scaling of successive categories was considered as a tool for the detection of response styles. By means of a simulation study, we showed that this mathematically simple and straightforward approach appears to pick up response styles quite well. Even when the original preference structure is similar to a response style, dual scaling can be used to detect a response style provided that the amount of noise in the sample can be assumed to be sufficiently large.

Table 8: Relative Inter-Boundary Deviations: Midpoint Preferences

	α=1/4		α=1/3		α=1/2	
Boundaries	Mean	Std	Mean	Std	Mean	Std
1-2	-0.16	0.04	-0.09	0.04	-0.04	0.05
2-3	0.07	0.05	0.04	0.05	0.01	0.05
3-4	0.16	0.05	0.09	0.05	0.04	0.05
4-5	0.07	0.05	0.04	0.05	0.02	0.05
5-6	-0.15	0.04	-0.08	0.04	-0.03	0.04

As with any study, our study suffers from some limitations. First of all, in our simulation study we considered a design with 7 rating categories and 5 objects. It remains to be seen what the changes will be if one or both of these parameters are varied. Secondly, one may consider different response functions from the ones used in the simulation study. Thirdly, one may consider different ways to model the noise in our model. Perhaps different distributions lead to different outcomes. Important with respect to this is of course the question how the noise can be generated in a realistic fashion. Perhaps other simulation studies, used for example in the context of latent class analysis, can give some insights into this matter. Fourthly, related to the previous point, it may be interesting to compare the dual scaling method with respect to other methods that can be applied in this context. These limitations, and probably others that are not mentioned here, should be addressed in future work on this topic.

The existence of response styles in an important measurement problem. Especially in the context of international research the problem gains in urgency. In this paper, we addressed this issue by considering dual scaling as a detection tool for the response styles. The results of this novel application of dual scaling of successive categories are promising and indicate that dual scaling may indeed serve as a tool to to gauge the existence of response styles.

References

Baumgarter, H. & Steenkamp J. B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research, 38*, 143–156

Benzécri, J. P. (1973). *L'analyse des correspondances : (introduction, thorie, applications diverses, notamment l'analyse des questionnaires, programmes de calcul).* Paris: Dunod.

Greenacre, M. J. (1984). *Theory and Applications of Correspondence Analysis.* London: Academic Press

Greenleaf, E. A. (1992). Improving rating scale measures by detecting and correcting bias components in some response styles. *Journal of Marketing Research, 29*, 176–188.

Marin, G. Gamba, R. J.& Marin, B. V. (1992). Extreme response style and acquiescence among Hispanics: the role of acculturation and education. *Journal of Cross Cultural Psychology, 4*, 498–509.

Nishisato, S. (1980). *Analysis of Categorical Data: Dual Scaling and its Applications.* Toronto: University of Toronto Press

Nishisato, S. (1994). *Elements of Dual Scaling: An introduction to practical data analysis.* Hillsdale, New Jersey: Lawrence Erlbaum Associates

Nishisato, S.& Shue, W. J. (1984). A note on dual scaling of successive categories data. *Psychometrika, 49*, 493–500.

Torres, A.& Greenacre, M. J. (2002). Dual scaling and correspondence analysis of preferences, paired comparisons and ratings. *International Journal of Research in Marketing, 19*, 401–405.

Van de Velden, M. (2003). Optimal scaling of paired comparison data. *Journal of Classification, 21*, 89–109.

Van de Velden, M. (2000). Dual scaling and correspondence analysis of rank order data. In R. D. H. Heijmans, D. S. G. Pollock, & A., Satorra (Eds.), *Innovations in multivariate statistical analysis.* Dordrecht: Kluwer Academic Publishers pp. 87–99.

Van Herk, H. Poortinga, Y. H.,& Verhallen, T. M. (2004). Response styles in rating scales: evidence of method bias in data from six EU countries. *Journal of Cross Cultural Psychology, 35*, 346-360.