HOW TO USE THE BLACK BOX

by

Keith T. Poole
Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh, PA 15213

3 August 1998

Abstract

This paper is meant as a supplement to my August, 1998 ***AJPS*** article "Recovering a Basic Space From a Set of Issue Scales." I promised the editor and the reviewers of my article that I would provide *and support* the computer programs used in the article. Accordingly, this paper is part of a package of material that contains the FORTRAN code and the executables of the programs used in the article.

1. Introduction

This paper is meant as a supplement to my ***AJPS*** article "Recovering a Basic Space From a Set of Issue Scales. The body of the paper shows material deleted from the original paper to conserve journal space as well as additional empirical examples. Appendices A, B, and C show researchers how to use the various computer programs that implement the model shown in the article.

Section 2 shows how the method I develop in my ***AJPS*** article is related to Factor Analysis. Section 3 reports some Monte-Carlo work that was cut out of the final version of the paper. Section 4 shows the relationship between the method and Aldrich-McKelvey scaling (1977). In effect, the method can be used to perform an Aldrich-McKelvey scaling of an issue scale in more than one dimension. Finally, Section 5 shows some additional empirical applications.

2. Relationship With Standard Factor Analysis

A standard method of analyzing a rectangular data matrix is to compute a correlation matrix between variables (the columns of the data matrix) and then factor analyze (principal components or maximum likelihood) the correlation matrix. Factor analysis has its own special nomenclature and the method of its presentation varies from author to author. However, in its simplest form, principal components, it is simply eigenvalue/eigenvector decomposition and its connection to singular value decomposition is easily shown.

For example, suppose the n by m data matrix, **X**, has no missing data and is standardized such that each column sums to zero and the sum of the squared entries of the column sums to one. Note that this is the transformation

$$\mathbf{x}_{ij} = \frac{(\mathbf{x}^*_{ij} - \bar{\mathbf{x}}_j)}{\mathbf{n}^{1/2}\mathbf{s}_j} \tag{1}$$

where $\mathbf{x}^*_{ij}$ is the original matrix entry, $\bar{\mathbf{x}}_j$ is the mean of the jth column, and $\mathbf{s}_j$ is the standard deviation of the jth column. Given the transformation given in equation (1), the Pearson correlation matrix is simply **R = X'X.**

Alternatively, most authors (e.g., Harman, 1970; Van de Geer, 1971) assume that **X** is in standard deviation form. That is:

$$\mathbf{x}_{ij} = \frac{(\mathbf{x}^*_{ij} - \bar{\mathbf{x}}_j)}{\mathbf{s}_j} \tag{2}$$

Written in this form the Pearson correlation matrix is $\mathbf{R} = \frac{\mathbf{X'X}}{\mathbf{n}}$. This approach has the awkward result of having the 1/n in various equations. Accordingly, I use the simpler approach of equation (1). This has no material effect on the discussion below except to simplify the expressions.

Let the singular value decomposition of **X** be $\mathbf{ULV'}$, where **U** is an n by m matrix such that $\mathbf{U'U} = \mathbf{I}_m$, $\mathbf{L}^{1/2}$ is a m by m diagonal matrix of singular values, and **V** is a m by m matrix such that $\mathbf{V'V} = \mathbf{I}_m$.

To perform a principal components analysis compute the correlation matrix,

$$\mathbf{R} = \mathbf{X'X} = \mathbf{V}\mathbf{L}^2\mathbf{V'} \qquad \textbf{(2)}$$

and then perform a standard eigenvalue/eigenvector decomposition of **R**. Note that the eigenvalues of **R** are the squared singular values of **X**. The **factor matrix** is the m by m matrix $\mathbf{VL}$ and the **factor scores** are the n by m matrix **U**, where **U** is from the singular value decomposition of **X**. Using the terminology of Harman (1970):

$$\mathbf{F} = \mathbf{U} \quad \text{and} \quad \mathbf{A} = \mathbf{VL}$$

In terms of the model I use in the ***AJPS*** article,

$$\mathbf{X}_0 = [\mathbf{Y}\,\mathbf{W}' + \mathbf{J}_n\underline{\mathbf{c}}']_0 + \mathbf{E}_0$$

let $\underline{\mathbf{c}}' = \underline{\mathbf{0}}$, no missing data, and ignoring the error term for the moment, then

$$\mathbf{Y} = \mathbf{U}\mathbf{L}^{1/2} = \mathbf{F}\mathbf{L}^{1/2} \text{ and } \mathbf{W} = \mathbf{V}\mathbf{L}^{1/2} = \mathbf{A}\mathbf{L}^{-1/2} \qquad \textbf{(3)}$$

so that **Y** and **W** are simply related to the factor scores and the factor matrix, respectively.

Indeed, even when **X** is in the form of the ***AJPS*** paper, $[\mathbf{Y}\,\mathbf{W}' + \mathbf{J}_n\underline{\mathbf{c}}']_0 + \mathbf{E}_0$, the estimated **W** matrix, $\hat{\mathbf{W}}$, will be highly correlated with the factor matrix, **A.** For example, Table 1 shows the SPSS output for a principal components analysis of the 1980 issue scale example shown in the ***AJPS*** paper. The data set was read into SPSS and the correlation matrix was computed using the pair-wise deletion option. The first part of the table shows the eigenvalue table and the second part of the table shows the Factor Matrix, **A**, labeled "Component Matrix" in the SPSS output.

The r-squares between the 3 columns of the Factor Matrix shown in Table 1 and the columns of the $\hat{\mathbf{W}}$ shown in Table 4 of the ***AJPS*** article are .929, .802, and .223 respectively. In other words, the first two dimensions are essentially the same. This makes sense because I found that only two of the fourteen $\hat{\mathbf{w}}$'s for the 3rd dimension to be statistically significant whereas seven of the fourteen $\hat{\mathbf{w}}$'s for the 2nd dimension and all for the 1st dimension were statistically significant. The individual placements appear to be at most two-dimensional.

Further evidence of the data being at most two-dimensional are the scree plots shown in Figure 1. The dotted line in the upper plot shows the eigenvalues of the correlation matrix from the SPSS output. The solid line in the upper plot shows the squared singular values for the data transformed as in equation (1). That is, let $\mathbf{X^* = X_0}$ for the non-missing entries and let $\mathbf{X^* = Y\,W'}$ (using three dimensions) for the missing entries. Column means were computed using the non-missing entries of $\mathbf{X_0}$**,** and each entry of $\mathbf{X^*}$ was transformed as shown in equation (1). The singular values of this matrix were then squared to make them comparable to the eigenvalues extracted from the correlation matrix. These two series are virtually identical with the values of the eigenvalues/squared singular values falling off fairly smoothly from the elbow at the 3rd value through the 14th value. This is a clear indication that the data are most likely to be two-dimensional.

Table 1
SPSS Output for 1980 Issue Scale Example

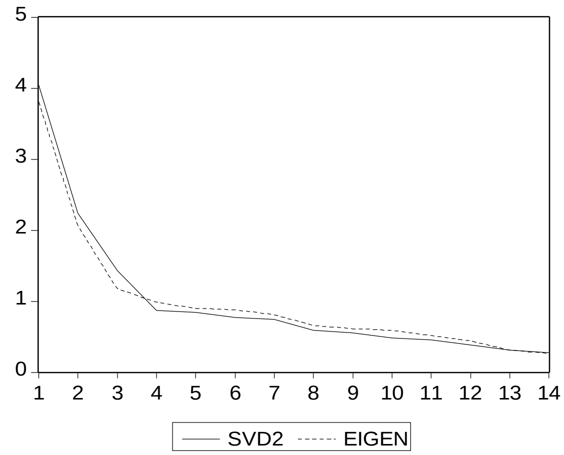
Extraction Method: Principal Component Analysis.
Total Variance Explained

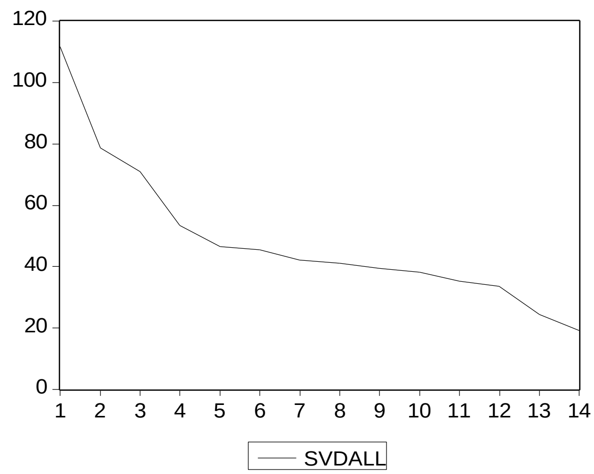
	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.432	31.654	31.654	4.432	31.654	31.654
2	2.055	14.677	46.331	2.055	14.677	46.331
3	1.155	8.252	54.583	1.155	8.252	54.583
4	.916	6.546	61.129			
5	.879	6.281	67.410			
6	.790	5.644	73.054			
7	.780	5.570	78.625			
8	.630	4.497	83.122			
9	.599	4.278	87.399			
10	.541	3.865	91.264			
11	.409	2.919	94.183			
12	.394	2.812	96.995			
13	.223	1.595	98.590			
14	.197	1.410	100.000			

Extraction Method: Principal Component Analysis.

Figure 1

Eigenvalues vs. Singular Values





The lower plot shows the singular values of **X*** minus the actual column means, that is, $\mathbf{X}^{*} - \mathbf{J}_{n}\underline{\mathbf{c}}'$, where $\underline{\mathbf{c}}$ is the vector of actual column means. The column means are subtracted from **X*** because if the means are large numerically – as they are for issue scales – then the first singular value of **X*** will be very large compared to the others. This results from the fact that the sum of the squared singular values is equal to the sum of the squared values of the matrix. That is:

$$\sum_{j=1}^{m} \lambda^{2} = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}^{2}$$

Here the singular values fall off a little less smoothly. The decline from the 2nd to the third value is not as dramatic as it is for the eigenvalues of the correlation matrix, but the plot clearly indicates that the data is *at most* three-dimensional (see the output for the 1980 issue scale example in Appendix A). In this context, a singular value decomposition of **X** yields essentially the same information as a principal components analysis of the correlation matrix.

3. Additional Monte Carlo Tests of the Model

This subsection was deleted from the final version of the paper in order to conserve journal space. The purpose of these tests is to show the ability of the procedure to estimate the Eckart-Young lower rank approximation matrix of an arbitrary matrix of real numbers with missing entries. The table and equation references are to those in the **AJPS** paper. The table deleted along with this subsection was the original Table 4. Here I will refer to it as 4* to avoid confusion with the **AJPS** article.

Estimating Eckart-Young Approximation Matrices

The Monte Carlo results reported in Tables 2 and 3 show that the procedure does an excellent job of estimating **Y** , **W**, and **c̲** when the observed data are in the form shown in equation (1). The purpose of this subsection is to show that the procedure also can be used as a general-purpose tool to obtain the Eckart-Young approximation matrix $\mathbf{U}\mathbf{L}_s\mathbf{V}'$ of any rectangular matrix of real numbers with missing entries.

In this application $\hat{\mathbf{y}}\hat{\mathbf{W}}' \ \mathbf{J}_n\underline{\hat{\mathbf{c}}}'$ must be used with some caution. Recall that if a matrix **X** is of rank s, then subtracting off the column means, $\mathbf{X} - \mathbf{J}_n\underline{\mathbf{c}}'$, in most circumstances, does not change the rank.[1] However, the converse is not necessarily true. By construction, $\hat{\mathbf{y}}\hat{\mathbf{W}}'$ has rank s and the columns of $\hat{\mathbf{y}}\hat{\mathbf{W}}'$ sum to zero. Adding a vector of constants, $\hat{\mathbf{y}}\hat{\mathbf{W}}' \ \mathbf{J}_n\underline{\hat{\mathbf{c}}}'$, will usually increase the rank to s+1. However, the first s singular values of $?\hat{\mathbf{W}}' \ \mathbf{J}_n\underline{\hat{\mathbf{c}}}'$ will be quite large in comparison to the $s+1^{st}$ singular value. If the observed data are in form given by equation (1), then the closer $\hat{\mathbf{y}}\hat{\mathbf{W}}'$ is to the true **Y W'** (which, by construction, is $\mathbf{X} - \mathbf{J}_n\underline{\mathbf{c}}'$), the smaller the $s+1^{st}$ singular value. In addition, if the column means are not of interest, then the Monte Carlo results in Table 2 show that $\hat{\mathbf{y}}\hat{\mathbf{W}}'$ is an excellent approximation of the true **Y W'** matrix even at substantial levels of error and missing data.

In order to test the ability of the procedure to estimate the Eckart-Young approximation matrix of an arbitrary matrix of real numbers of full rank, just the first s singular values of $\hat{\mathbf{y}}\hat{\mathbf{W}}' \ \mathbf{J}_n\underline{\hat{\mathbf{c}}}'$ were utilized. That is, if the singular value decomposition of

$\hat{\mathbf{y}}\ \hat{\mathbf{W}}$ ¢ $\mathbf{J}_n \underline{\hat{\mathbf{c}}}$ ¢ is $\mathbf{ULV'}$ where $\mathbf{L}$ is an s+1 by s+1 matrix, then $\mathbf{UL_sV'}$ was used as the approximation matrix.

Similar to the method used for Table 2, to construct **X** of full rank m, **U** and **V** were obtained from a singular value decomposition of an n by m matrix of uniform [-1,1] random numbers. The first three singular values of **X** were set so that when the column means are subtracted from **X**, the singular values of the resulting matrix, $\mathbf{Y W'}$, were approximately 50, 35, 20, 5, 5, …, 5. This was done so as to approximate a situation where the first three dimensions largely account for the structure of the matrix but the matrix is of full rank.

Missing data was created in the same fashion as described for Table 2.

In these tests there is no error process so the only difference between trials is the pattern of missing data. Each entry in Table 4* is the average of 10 trials. The standard deviations are shown in parentheses.

Table 4* About Here

The first three columns of Table 4 show the number of rows, the number of columns, and the rank of the approximation matrix**.** The fourth column, r-square with Eckart-Young, shows the average squared Pearson correlation between the nm elements in true Eckart-Young matrix and the reproduced Eckart-Young matrix. The fifth column shows the average squared Pearson correlation between the true basic dimensions and the estimated basic dimensions corresponding to the rank of the approximation; that is, the average of the r-squares computed between each column of the true **Y** matrix, and its corresponding column in $\hat{\mathbf{Y}}$.[2] Finally, the sixth column shows the percentage of missing entries.

Table 4* shows that the procedure does a good job estimating lower rank approximations when a substantial portion of the matrix is missing. Not surprisingly, the lower the level of missing data and the larger the matrix, the better the approximation. With 25 percent missing data the r-squares all exceed .97. Even with 70 percent missing data the procedure will do a reasonable job if the size of the matrix is large enough.

4. Empirical Application Deleted From AJPS Article

This application was deleted from the final draft of the **AJPS** article to conserve journal space. In this application the scaling method is applied to a *transposed* matrix in which the number of columns *is much larger* than the number of rows. What I show below is that the general method developed in the **AJPS** article can be used to perform Aldrich-McKelvey scaling of an issue scale in more than one dimension.

Aldrich and McKelvey (1977) in effect solved the Likert scale problem. In my opinion, the Aldrich and McKelvey paper is the most under appreciated achievement in political methodology. In part this is due to the fact that, as I note below, the standard error of the estimate of their model is clearly biased downwards. However, as I have argued elsewhere (Palfrey and Poole, 1987), *this is an advantage not a defect*! Namely, the Aldrich-McKelvey scaling method can be used as a powerful filter. Respondents who see the political universe as *backward* (namely, Reagan to the left of Carter), clearly have a very low level of information about politics.

The big advantage of applications like the one shown below is that a researcher can check to see if the scale *really is one-dimensional*. That is, scales with labeled endpoints are designed to be one-dimensional. Performing the decomposition shown below can roughly test this. Namely, the r-square in one dimension should be very large and the increment to adding a second dimension should be quite small.

Note that because the number of political stimuli being placed on an issue scale is usually not too large, this application should be done with some caution. For the 1980 scale shown below, there are only 6 stimuli. Consequently, if respondents are included that placed only 4 of the 6 stimuli, then if 3 dimensions are estimated the r-square for these respondents will be 1.0! Because of this I only estimated 2 dimensions with two missing responses (see Appendix C for the output files).

The one-dimensional fit of the model was an r-square of .7541 and a standard error of the estimate of .9843. In two dimensions the r-square was .8645 and the standard error of the estimate was .8447. These fits plus the estimated configuration shown in Table 6* show, in my opinion, that the scale is indeed one-dimensional and that the second dimension is largely capturing respondent confusion about where to place Anderson on the scale.

Note that in Table 6* I show the coordinates as $\hat{\mathbf{Y}}$. In the computer code I simply write out the singular vectors to make it easier to compare with the Aldrich-McKelvey coordinates (which is an eigenvector).

Analysis of 1980 Post-Election Liberal-Conservative Scale

The purpose of this application is to show the connection between the procedure developed here and the scaling method developed by Aldrich and McKelvey (1977) to analyze seven-point scales. The Aldrich-McKelvey scaling method is a one dimensional version of the model expressed in equation (1) – that is, the original model as expressed in equation (1B) applied to a transposed matrix where m > n. In this application I will analyze the responses to the 1980 Post-Election Liberal-Conservative seven-point scale. This is one of the fourteen scales analyzed in the previous subsection.

In the Aldrich-McKelvey framework, the matrix **X** is an n by m matrix where the rows are the respondents' perceived positions of the m stimuli on the scale.
The model they estimate is

$$\underline{\mathbf{y}}\mathbf{J}_{\mathbf{n}}' + \mathbf{E} = \mathbf{X}\mathbf{W} + \mathbf{J}_{\mathbf{m}}\underline{\mathbf{c}}' \qquad (17)$$

where $\underline{\mathbf{y}}$ is an m length vector of underlying stimulus coordinates, **W** is a n by n diagonal matrix of weights, $\mathbf{J}_{\mathbf{m}}$ is an m length vector of ones, $\underline{\mathbf{c}}$ is an n length vector of constants, and **E** is an m by n matrix of error terms. Aldrich and McKelvey assume that the respondents correctly perceive the true underlying configuration subject to some random perceptual error, **E**, and report a linear transformation of that true configuration. Their scaling method estimates $\underline{\mathbf{y}}$ using **X**, and **W** and $\underline{\mathbf{c}}$ are estimated using $\hat{\underline{\mathbf{y}}}$ and **X** with ordinary least squares.

The Aldrich-McKelvey scaling method is, in effect, an s = 1 version of equation (1). Solving for **X** in (17) produces

$$\mathbf{X}' = \underline{\mathbf{y}}\mathbf{W}^* + \mathbf{J}_{\mathbf{m}}\underline{\mathbf{c}}^* + \mathbf{E}^* \qquad (18)$$

where $\mathbf{W^*} = \mathbf{W}^{-1}\mathbf{J_n}$, $\underline{\mathbf{c}}^* = -\mathbf{W}^{-1}\underline{\mathbf{c}}$, and $\mathbf{E^*} = \mathbf{EW}^{-1}$. Equation (18) is identical to equation (1) the only difference being the reversal of the roles of n and m.

Aldrich and McKelvey require that $\underline{\mathbf{y}}'\underline{\mathbf{y}} = 1$ and missing entries are not allowed in **X**. The procedure outlined in Section 2 based on the model stated in equation (1) can be regarded as a generalization of the Aldrich-McKelvey scaling procedure to more than one dimension.

In this application the rows of the data set are the political stimuli and the columns are the respondents' perceptions of where on the seven-point scale the stimuli are. Consequently, there are n political stimuli (note the reversal of role of n from the equations above) and the basic space coordinates of the political stimuli are given in $\hat{\mathbf{Y}}$. Recall that, by equation (1B), if **X** has no missing entries and no error, then it has rank s. However, because m > n, subtracting off the column means reduces the rank of **X** by one provided that the columns do not already sum to zero. In this case the number of basic dimension is s-1 so that $\hat{\mathbf{Y}}$ is an n by s-1 matrix, $\hat{\mathbf{W}}$ is m by s-1 and $\hat{\underline{\mathbf{c}}}$ is an m length vector where $\hat{\mathbf{W}}$ and $\hat{\underline{\mathbf{c}}}$ are the linear mappings for the m respondents.

Table 6* shows $\hat{\mathbf{Y}}$ for two basic dimensions along with the corresponding one dimensional vector estimated by the Aldrich-McKelvey procedure. The first basic dimension is the liberal-conservative dimension and the order of the political stimuli – from Ted Kennedy at the far left to John Anderson near the center of the spectrum to Ronald Reagan at the far right – is intuitively appealing. The second basic dimension essentially separates John Anderson from everyone else. The standard errors were computed using a bootstrap procedure identical to that described earlier except now the columns (respondents) are being sampled with replacement. The standard errors are based on 100 trials.

Table 6* About Here

These standard errors must be taken with a grain of salt, however, because, as Aldrich and McKelvey (1977) note, a respondent "…who sees things backwards … contributes to a better fit to the 'true' space" (p. 116). That is, respondents who perceive a mirror image of the true configuration improve the fit of the model so that the standard errors in Table 6* underestimate the true standard errors. However, Monte Carlo work done by Aldrich and McKelvey and Palfrey and Poole (1987), show that the recovery of the stimulus configuration is robust to violations of the error assumptions and is very accurate even when the error level is very high and a large number of respondents are reporting mirror or semi-mirror images.

The fourth column of Table 6* shows the first basic dimension normalized so that it can be directly compared to the Aldrich-McKelvey configuration shown in the fifth column. The two configurations are, not surprisingly, virtually identical. The differences are due to the slightly different samples analyzed by the two procedures. Of the 888 respondents used to estimate the two basic dimensions, 643 had no missing data and were used in the Aldrich-McKelvey procedure.

5. **Additional Empirical Examples**

a. Recovering a Basic Space From the 1992 Issue Scales

Table 2 shows an analysis of fifteen issue scales from the 1992 NES survey. The survey included a panel as well as a cross section with some respondents included in both groups. The table is laid out the same as Table 4 in the ***AJPS*** article.

Table 2 About Here

The results are similar to those for the 1980 scales. The first dimension is clearly liberal/conservative and the second dimension appears to be picking up the abortion and women's equal role questions. However, in contrast to the 1980 results, the first dimension only accounts for about forty percent of the variance (r-square of .394 versus .512 for 1980) and the second dimension is not as strongly related to abortion and women's rights as it was in 1980. If conventional statistical tests were applied using the bootstrapped standard errors, then all the $\hat{\mathbf{w}}$'s for the first dimension, fourteen of the fifteen for the second, and eight of the fifteen for the third are statistically significant.

The data are clearly *at most* three-dimensional. Adding a fourth dimension improves the overall r-square to .674 but *only one* of the estimated $\hat{\mathbf{w}}$'s is statistically significant.

Because the set of scales is not the same between 1980 and 1992 it is not possible to make claims about changes in the structure and the fit of the basic space over time. That requires a panel and would be an interesting topic for future research. However, it certainly appears that the basic space *is low dimensional*. It appears that only two basic dimensions – one capturing general liberalism/conservatism and one picking up issues related to the rights of women – give a good summary of mass attitudes.

References

Aldrich, John and Richard D. McKelvey. 1977. " A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections." *American Political Science Review*, 71:111-130

Eckart, Carl and Gale Young. 1936. "The Approximation of One Matrix By Another of Lower Rank." *Psychometrika*, 1: 211-218.

Harman, Harry H. 1970. *Modern Factor Analysis*. Chicago: University of Chicago Press.

Palfrey, Thomas R. and Keith T. Poole. 1987. "The Relationship Between Information, Ideology, and Voting Behavior." *American Journal of Political Science*, 31:511-530.

Schonemann, Peter H. 1966. " A Generalized Solution of the Orthogonal Procrustes Problem." *Psychometrika*, 31:1-10.

Schonemann, Peter H. and R.M. Carroll. 1970. "Fitting One Matrix to Another Under Choice of a Central Dilation and Rigid Motion." *Psychometrika*, 35:245-256

Van de Geer, John P. 1971. *Introduction to Multivariate Analysis for the Social Sciences*. San Francisco: W.H. Freeman and Company.

Table 4*
Monte Carlo Tests of Eckart-Young Lower Rank Approximation

(Singular Values of Y W': 50, 35, 20, 5, 5, 5, …., 5)

N	M	Rank of Approximation	R^2 With E-Young	R^2 With True Y	Percent Missing
100	25	1	.932 (.015)	.949 (.016)	50
100	25	2	.954 (.007)	.959 (.012)	50
100	25	3	.961 (.004)	.947 (.017)	50
500	25	1	.962 (.005)	.966 (.004)	50
500	25	2	.966 (.003)	.965 (.003)	50
500	25	3	.970 (.001)	.953 (.004)	50
1000	25	1	.822 (.032)	.828 (.031)	70
1000	25	2	.845 (.019)	.843 (.020)	70
1000	25	3	.915 (.006)	.873 (.022)	70
100	50	3	.989 (.001)	.986 (.004)	25
250	50	2	.992 (.000)	.993 (.001)	25
500	15	2	.987 (.001)	.986 (.003)	25

Table 6*
1980 Liberal-Conservative Scale

Political Stimulus	First Basic Dimension	Second Basic Dimension	Normalized First Basic Dimension	Aldrich-McKelvey
Jimmy Carter	-2.234 (0.088)	-3.203 (0.301)	-0.229	-0.232
Ronald Reagan	5.685 (0.059)	-0.773 (0.298)	0.582	0.582
Ted Kennedy	-4.703 (0.091)	0.007 (0.539)	-0.482	-0.485
John Anderson	-0.757 (0.099)	6.756 (0.833)	-0.078	-0.066
Republican Party	5.089 (0.077)	-0.759 (0.137)	0.521	0.517
Democratic Party	-3.080 (0.095)	-2.029 (0.277)	-0.315	-0.317

Table 2A
Overall Fit Statistics for Fifteen 1992 Issue Scales

S	N	% Missing	R^2	Standard Error of Estimate	Singular Values of ŷ Ŵ ¢
1	1264	17.6	.394	1.392	113.526
2	1264	17.6	.519	1.300	86.803
3	1264	17.6	.604	1.242	76.235

Table 2B
Fit Statistics by Issue
(Bootstrapped Standard Errors in Parentheses)

Issue	n_j	$\hat{c}_j$	$\hat{w}_1$	$\hat{w}_2$	$\hat{w}_3$	r-square 1	r-square 2	r-square 3
Liberal/Conservative	924	4.19	-2.56	0.82	-0.79	.301	.327	.372
		(0.04)	(0.16)	(0.42)	(0.89)			
Women's Equal Role	606	2.51	-3.25	4.32	-3.29	.188	.553	.749
		(0.07)	(0.24)	(0.39)	(1.09)			
Defense Spending	1158	3.59	-2.48	1.54	3.09	.202	.385	.529
		(0.05)	(0.16)	(0.32)	(0.75)			
Government Jobs	1143	4.17	-3.48	-4.09	-1.23	.419	.650	.661
		(0.06)	(0.17)	(0.36)	(0.86)			
Govt Help Minorities	1191	4.51	-3.44	-2.38	0.78	.420	.457	.489
		(0.06)	(0.19)	(0.45)	(1.15)			
Govt Provide Services	1122	4.36	1.98	2.18	3.22	.179	.319	.478
		(0.05)	(0.16)	(0.43)	(1.08)			
Abortion	1239	2.87	0.97	-1.94	0.66	.027	.215	.306
		(0.03)	(0.11)	(0.21)	(0.39)			
Liberal/Conservative	627	4.44	-2.83	1.45	-0.48	.328	.382	.412
		(0.07)	(0.23)	(0.60)	(1.27)			
Defense Spending	846	4.01	-2.87	1.37	3.37	.269	.442	.584
		(0.06)	(0.16)	(0.35)	(0.69)			
Govt Provide Services	1063	4.17	2.35	2.06	2.46	.243	.374	.447
		(0.05)	(0.19)	(0.42)	(1.02)			
Defense Spending	1135	3.53	-2.29	1.28	2.90	.204	.371	.529
		(0.04)	(0.13)	(0.21)	(0.53)			
Government Jobs	1141	4.29	-3.12	-3.02	-0.13	.380	.506	.512
		(0.05)	(0.17)	(0.31)	(0.82)			
Abortion	1233	2.95	1.22	-1.91	0.76	.061	.247	.348
		(0.03)	(0.11)	(0.20)	(0.36)			
Urban Unrest	972	3.35	-3.88	-0.91	3.22	.383	.398	.526
		(0.07)	(0.24)	(0.81)	(2.30)			
Women's Equal Role	1225	2.32	-2.89	3.29	-2.30	.160	.399	.631
		(0.05)	(0.17)	(0.42)	(0.88)			

Appendix A: BLACKBOX.FOR

1. Introduction

BLACKBOX.FOR is a FORTRAN program that implements the model discussed in "Recovering a Basic Space From a Set of Issue Scales." The program reads a "control card" file (BTSTR.DAT) and the data file (usually an NES dataset) and writes three output files: BLACK23.DAT, BLACK24.DAT, and BLACK28.DAT. BLACK23.DAT is a file that contains various information about the estimation, BLACK24.DAT is the output file for the respondent parameters (the p by s **Y** matrix for k=1,2,…,s), and BLACK28.DAT is the output file for the issue scale parameters (the n by s **W** matrix and the vector of constants, **c**).

The program has been compiled for both the Pentium P5 processor as well as the Pentium P6 (Pentium II) processor. These executables are BLACK5.EXE and BLACK6.EXE respectively. They will run under both Windows 95 and Windows NT.

The example below is from the ***AJPS*** article and it uses the 14 issue scales from 1980 NES cross-sectional survey data set -- NES1980.DAT.

2. Input File: BTSTR.DAT

The first line of the input file (see next page) gives the name of the data set being analyzed. In this case, the 1980 NES data is in the subdirectory \NES. That is, the program, BLACK6.EXE is in the root directory. Note that if NES1980.DAT could be placed in the same directory as the program the first line would simply be NES1980.DAT.

The second line of the input file is the title of the scaling

The third line contains, in order, the number of basic dimensions to be estimated, the number of issue scales, the maximum number of missing data values for the issue scales, and the minimum number of responses for a respondent to be included in the analysis. Note that this is fixed format, namely, in FORTRAN syntax, 4I5. Hence, if you change any of the numbers be sure to ***not change the spacing***. For example, if you want 2 basic dimensions and there are 10 stimuli, 11 missing data values, and a minimum number of 5 responses, this line would be:

```
    2   10   11    5
```

The fourth line is the format statement for the 1968 NES data file. The I4 is the respondent ID number and the I1's (there are 14 of them) are the respondents' self-reported positions on the 14 issue scales. The "X"s indicate spaces in FORTRAN format syntax. This format statement can be figured out by using the ICPSR codebook for the election study. If you have problems figuring out how to do this, just send me E-Mail at kp2a@andrew.cmu.edu.

The remaining 28 lines contain the title of each issue scale and the missing value codes for the scale. For the 7-Point scales these are 0, 8, and 9 but for the two abortion scales the missing values are 0, 7, 8, 9. Note that these numbers are also fixed format.

```
 BLACKB\NES\NES1980.DAT
 DECOMPOSITION OF 14 1980 7-POINT SCALES
    3   14    4    8
(8X,I4,527X,I1,13X,I1,11X,I1,11X,I1,11X,I1,13X,I1,1217X,I1,36X,I1,17X,I1,
17X,I1,17X,I1,18X,I1,5X,I1,2X,I1)
LIBERAL/CONSERVATIVE
    0    8    9
DEFENSE
    0    8    9
GOVT SERVICES
    0    8    9
INFLATION
    0    8    9
ABORTION
    0    7    8    9
TAX CUTS
    0    8    9
LIBERAL/CONSERVATIVE
    0    8    9
GOVT HELP MINORITIES
    0    8    9
RUSSIA
    0    8    9
WOMENS EQUAL ROLE
    0    8    9
GOVT JOBS
    0    8    9
EQUAL RIGHTS AMEND
    0    8    9
BUSING
    0    8    9
ABORTION
    0    7    8    9
```

3. Output files for 1980 Issue Scale Example Shown in AJPS Article

a. BLACK23.DAT File

This file is the primary output file for the program. For ease of exposition I will annotate this file for the convenience of the reader. My comments will be preceded by #### signs.

The first part of the output file just echoes the lines in BTSTR.DAT. This is convenient because you can glance at this to make sure the starting file is configured and being read correctly.

```
\BLACKB\NES\NES1980.DAT
 DECOMPOSITION OF 14 1980 7-POINT SCALEs
    3   14    4    8
(8X,I4,527X,I1,13X,I1,11X,I1,11X,I1,11X,I1,13X,I1,1217X,I1,36X,I1,17X,I1,17X,I1,17X,I1,18X,
I1,5X,I1,2X,I1)
LIBERAL/CO
    0    8    9    0
DEFENSE
    0    8    9    0
GOVT SERVI
    0    8    9    0
INFLATION
    0    8    9    0
ABORTION
    0    7    8    9
TAX CUTS
    0    8    9    0
LIBERAL/CO
    0    8    9    0
GOVT HELP
    0    8    9    0
RUSSIA
    0    8    9    0
WOMENS EQU
    0    8    9    0
GOVT JOBS
    0    8    9    0
EQUAL RIGH
    0    8    9    0
BUSING
    0    8    9    0
ABORTION
    0    7    8    9
####
#### Here 1270 respondents have been included in the analysis.  The
#### program requires that a respondent place himself/herself on at
#### least 8 (see input file) scales.
####
 NUMBER OF CASES  1270
 ******************************************************************************
 ******************************************************************************
 ******************************************************************************

#### The program first estimates a one dimensional model, then two
#### dimensions, etc.

 NUMBER OF DIMENSIONS=   1
 ************************************************************************
```

This information is only given once.
The matix is 1270 by 14 and contains 17,780 entries. There
are 2509 missing entries or [2509/(14*1270)]*100 = 14.11% missing
entries. The Sum of Squares is computed around the grand mean of
the matrix. Hence, it is the sum of the squared differences between
the 15,271 non-missing entries and the matrix mean.

```
 NUMBER OF ROWS                =  1270
 NUMBER OF COLUMNS             =    14
 TOTAL NUMBER OF DATA ENTRIES = 15271
 NUMBER MISSING ENTRIES        =  2509
 PERCENT MISSING DATA          =       14.11136
 SUM OF SQUARES GRAND MEAN     =    52705.13281
 ******************************************************************************
```

This is the iteration record for the first basic dimension.
REG1 estimates **W** and **c** and REG2 estimates **y** .

```
 DIMENSION=  1 TOTAL SSE REG1=     26094.3320
 DIMENSION=  1 TOTAL SSE REG2=     25818.1855
 DIMENSION=  1 TOTAL SSE REG1=     25742.5508
 DIMENSION=  1 TOTAL SSE REG2=     25718.7637
 DIMENSION=  1 TOTAL SSE REG1=     25710.2461
 DIMENSION=  1 TOTAL SSE REG2=     25706.7617
 DIMENSION=  1 TOTAL SSE REG1=     25705.1016
 DIMENSION=  1 TOTAL SSE REG2=     25704.3594
 DIMENSION=  1 TOTAL SSE REG1=     25704.0059
 DIMENSION=  1 TOTAL SSE REG2=     25703.9063
 DIMENSION=  1 TOTAL SSE REG1=     25703.7559
 DIMENSION=  1 TOTAL SSE REG2=     25703.6621
 DIMENSION=  1 TOTAL SSE REG1=     25703.6855
 DIMENSION=  1 TOTAL SSE REG2=     25703.5859
 DIMENSION=  1 TOTAL SSE REG1=     25703.6719
 DIMENSION=  1 TOTAL SSE REG2=     25703.5996
 DIMENSION=  1 TOTAL SSE REG1=     25703.6699
 DIMENSION=  1 TOTAL SSE REG2=     25703.5859
```

The singular values for the one dimensional estimation are reported
below. Only the first **s+3** singular values are shown.

```
 SINGULAR VALUES OF ESTIMATED MATRICES
 FIRST COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES
 SECOND COLUMN: REPRODUCED MATRIX -- PSI*W + Jc
 THIRD COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES MINUS THE ORIGINAL COLUMN
MEANS
 FOURTH COLUMN:  PSI*W
  1   545.737   544.149   111.761   111.757
  2    97.628    93.424    68.908     0.000
  3    68.464     0.000    57.681     0.000
  4    57.486     0.000    54.781     0.000
 ******************************************************************************
```

Below are the constraint checks discussed in the AJPS article --
namely, the sum of the columns of **y** must equal zero, and:
y 'y = W'W = L where **L** is the s by s diagonal matrix of the
singular values of **y W'** which is the least squares estimate of
$[\mathbf{X_0 .- J_n \underline{c}'}]$.

```
 CONSTRAINT CHECKS ON PSI AND W
     SUM OF COLUMNS OF PSI
          0.0000
     PSI-TRANSPOSE*PSI
     1  111.7571
     W-TRANSPOSE*W
     1  111.7571
```

Here $\mathbf{y W' + J_n \underline{c}'}$ is constructed and the r-square between the elements

of $\mathbf{y W' + J_n \underline{c}'}$ and the original data matrix, $\mathbf{X_0}$,is computed as a
check on the estimation.
####

```
R-SQUARE    15271  0.512
```

Similar to the above, as a check, a singular value decomposition
of the estimated matrix, $\mathbf{y W'}$, is performed. The rank of the
matrix is reported (here it is one) along with the singular values.

```
 R-SQUARE    15271  0.512
 RANK CHECK OF PSI*W    1
    1  111.7571
    2    0.0000
    3    0.0000
    4    0.0000
    5    0.0000
    6    0.0000
```

####
The program now estimates all the above for two dimensions --
s = 2.
####

```
 ******************************************************************************
 NUMBER OF DIMENSIONS=   2
 ******************************************************************************
 DIMENSION=  1 TOTAL SSE REG1=     26094.3320
 DIMENSION=  1 TOTAL SSE REG2=     25818.1855
 DIMENSION=  1 TOTAL SSE REG1=     25742.5508
 DIMENSION=  1 TOTAL SSE REG2=     25718.7637
 DIMENSION=  1 TOTAL SSE REG1=     25710.2461
 DIMENSION=  1 TOTAL SSE REG2=     25706.7617
 DIMENSION=  1 TOTAL SSE REG1=     25705.1016
 DIMENSION=  1 TOTAL SSE REG2=     25704.3594
 DIMENSION=  2 TOTAL SSE REG1=     21380.9219
 DIMENSION=  2 TOTAL SSE REG2=     21119.8066
 DIMENSION=  2 TOTAL SSE REG1=     20949.8340
 DIMENSION=  2 TOTAL SSE REG2=     20838.7109
 DIMENSION=  2 TOTAL SSE REG1=     20766.0664
 DIMENSION=  2 TOTAL SSE REG2=     20717.7871
 DIMENSION=  2 TOTAL SSE REG1=     20685.3633
 DIMENSION=  2 TOTAL SSE REG2=     20663.1270
 DIMENSION=  2 TOTAL SSE REG1=     20646.6641
 DIMENSION=  2 TOTAL SSE REG2=     20547.0156
 DIMENSION=  2 TOTAL SSE REG1=     20538.1191
 DIMENSION=  2 TOTAL SSE REG2=     20532.5488
 DIMENSION=  2 TOTAL SSE REG1=     20528.8223
 DIMENSION=  2 TOTAL SSE REG2=     20526.2266
 DIMENSION=  2 TOTAL SSE REG1=     20524.4395
 DIMENSION=  2 TOTAL SSE REG2=     20523.1094
 DIMENSION=  2 TOTAL SSE REG1=     20522.2148
 DIMENSION=  2 TOTAL SSE REG2=     20521.5879
```

The singular values for the two dimensional estimation are reported
below. Only the first **s+3** singular values are shown. Hence, there
are now 5 rows being printed out.

```
 SINGULAR VALUES OF ESTIMATED MATRICES
 FIRST COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES
 SECOND COLUMN: REPRODUCED MATRIX -- PSI*W + Jc
 THIRD COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES MINUS THE ORIGINAL COLUMN
MEANS
 FOURTH COLUMN:  PSI*W
  1   546.795   545.439   110.997   110.981
  2    97.860    95.370    77.571    77.559
  3    74.628    71.390    57.259     0.000
  4    57.217     0.000    54.304     0.000
  5    47.271     0.000    47.049     0.000
 ******************************************************************************
```

####
Below are the constraint checks discussed in the AJPS article --
namely, the sum of the columns of **y** must equal zero, and:
y 'y = W'W = L where **L** is the s by s diagonal matrix of the
singular values of **y W'** which is the least squares estimate of
[X_0.- $J_n\underline{c}$']. Now two by two matrices are being printed out.
####

```
 CONSTRAINT CHECKS ON PSI AND W
     SUM OF COLUMNS OF PSI
          0.0000    0.0000
     PSI-TRANSPOSE*PSI
     1  110.9815    0.0000
     2    0.0000   77.5591
     W-TRANSPOSE*W
     1  110.9814    0.0000
     2    0.0000   77.5591
 R-SQUARE    15271  0.611
 RANK CHECK OF PSI*W    2
    1  110.9815
    2   77.5591
    3    0.0000
    4    0.0000
    5    0.0000
    6    0.0000
```

####
Three dimensions -- the maximum set in the input file -- are
now estimated.
####

```
 ******************************************************************************
 NUMBER OF DIMENSIONS=   3
 ******************************************************************************
 DIMENSION=  1 TOTAL SSE REG1=     26094.3320
 DIMENSION=  1 TOTAL SSE REG2=     25818.1855
 DIMENSION=  1 TOTAL SSE REG1=     25742.5508
 DIMENSION=  1 TOTAL SSE REG2=     25718.7637
 DIMENSION=  1 TOTAL SSE REG1=     25710.2461
 DIMENSION=  1 TOTAL SSE REG2=     25706.7617
 DIMENSION=  1 TOTAL SSE REG1=     25705.1016
 DIMENSION=  1 TOTAL SSE REG2=     25704.3594
 DIMENSION=  2 TOTAL SSE REG1=     21380.9219
 DIMENSION=  2 TOTAL SSE REG2=     21119.8066
 DIMENSION=  2 TOTAL SSE REG1=     20949.8340
 DIMENSION=  2 TOTAL SSE REG2=     20838.7109
 DIMENSION=  2 TOTAL SSE REG1=     20766.0664
 DIMENSION=  2 TOTAL SSE REG2=     20717.7871
 DIMENSION=  2 TOTAL SSE REG1=     20685.3633
 DIMENSION=  2 TOTAL SSE REG2=     20663.1270
 DIMENSION=  3 TOTAL SSE REG1=     17706.8184
 DIMENSION=  3 TOTAL SSE REG2=     17619.4512
 DIMENSION=  3 TOTAL SSE REG1=     17548.3496
 DIMENSION=  3 TOTAL SSE REG2=     17489.4707
 DIMENSION=  3 TOTAL SSE REG1=     17440.7461
 DIMENSION=  3 TOTAL SSE REG2=     17400.1348
 DIMENSION=  3 TOTAL SSE REG1=     17365.9629
 DIMENSION=  3 TOTAL SSE REG2=     17336.9277
 DIMENSION=  3 TOTAL SSE REG1=     17296.1191
 DIMENSION=  3 TOTAL SSE REG2=     17058.3926
 DIMENSION=  3 TOTAL SSE REG1=     17023.2207
 DIMENSION=  3 TOTAL SSE REG2=     16995.1074
 DIMENSION=  3 TOTAL SSE REG1=     16971.8438
 DIMENSION=  3 TOTAL SSE REG2=     16952.2773
 DIMENSION=  3 TOTAL SSE REG1=     16935.7246
 DIMENSION=  3 TOTAL SSE REG2=     16921.6406
 DIMENSION=  3 TOTAL SSE REG1=     16909.6445
 DIMENSION=  3 TOTAL SSE REG2=     16899.3594
 SINGULAR VALUES OF ESTIMATED MATRICES
 FIRST COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES
 SECOND COLUMN: REPRODUCED MATRIX -- PSI*W + Jc
 THIRD COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES MINUS THE ORIGINAL COLUMN
MEANS
```

```
 FOURTH COLUMN:  PSI*W
  1   546.947   546.331   111.614   111.561
  2    99.310    98.669    78.557    78.521
  3    75.528    75.265    70.796    70.679
  4    63.612    52.554    53.402     0.000
  5    46.900     0.000    46.409     0.000
  6    46.021     0.000    45.545     0.000
 ******************************************************************************
```

####
Below are the constraint checks discussed in the AJPS article --
namely, the sum of the columns of **y** must equal zero, and:
y'y = W'W = L where **L** is the s by s diagonal matrix of the
singular values of **yW'** which is the least squares estimate of
[$\mathbf{X_0}$**.- $\mathbf{J_n}$c']**. Note that the matrices are now 3 by 3.
####

```
 CONSTRAINT CHECKS ON PSI AND W
     SUM OF COLUMNS OF PSI
          0.0000    0.0000    0.0000
     PSI-TRANSPOSE*PSI
     1  111.5610    0.0000    0.0000
     2    0.0000   78.5212    0.0000
     3    0.0000    0.0000   70.6788
     W-TRANSPOSE*W
     1  111.5609    0.0000    0.0000
     2    0.0000   78.5211    0.0000
     3    0.0000    0.0000   70.6788
 R-SQUARE    15271  0.679
 RANK CHECK OF PSI*W    3
    1  111.5610
    2   78.5211
    3   70.6788
    4    0.0000
    5    0.0000
    6    0.0000
 ****************************************************************************
```

####
Below is the estimation summary. Note that the error and the
explained always add to the total sum of squares. The standard
error of the estimate is also computed for each dimension.
The singular values above and the entries below are reported
in Table 4A of the **AJPS** article. The slight differences
between these numbers and those in the article are due to some
marginal improvements made in the program code.
####

```
 ITERATION RECORD
DIM     ERROR    EXPLAINED      PERCENT  CUM PERCENT   R-SQUARE  STD ERR EST
  1  25703.6367  27001.4961     51.2312     51.2312      0.5123  1.3549
  2  20521.4863  32183.6465      9.8323     61.0636      0.6106  1.2703
  3  16899.3535  35805.7813      6.8724     67.9360      0.6794  1.2158
 ****************************************************************************
 ****************************************************************************
```

b. BLACK24.DAT

Below is the output of the respondent parameters -- **Y** -- for one, two, and three dimensions. The first integer is a simple counter and runs, in this instance, from 1 to 1270 -- the number of respondents in the analysis. The second number is the actual case number in the NES1980.DAT file.

```
     1      1 -0.281
     2      2  0.595
     3      3  0.517
     4      5 -0.111
     5      7  0.398
     6      8  0.368
     7      9 -0.364
     8     10 -0.095
     9     11  0.090
    10     12  0.148
        etc.
        etc.
  1261   1756 -0.264
  1262   1757  0.024
  1263   1758 -0.158
  1264   1759 -0.419
  1265   1760 -0.324
  1266   1761 -0.011
  1267   1762 -0.332
  1268   1763  0.296
  1269   1764 -0.123
  1270   1765 -0.033
     1      1 -0.282 -0.390
     2      2  0.607 -0.255
     3      3  0.675 -0.375
     4      5 -0.099  0.170
     5      7  0.390  0.023
     6      8  0.383 -0.120
     7      9 -0.378  0.267
     8     10 -0.108  0.253
     9     11  0.072 -0.703
    10     12  0.145 -0.004
        etc.
        etc.
  1261   1756 -0.267  0.050
  1262   1757  0.023 -0.013
  1263   1758 -0.162  0.327
  1264   1759 -0.456  0.385
  1265   1760 -0.338  0.419
  1266   1761 -0.018  0.111
  1267   1762 -0.297 -0.142
  1268   1763  0.296 -0.264
  1269   1764 -0.129  0.320
  1270   1765 -0.054  0.189
     1      1 -0.281  0.392  0.024
     2      2  0.600  0.213  0.197
     3      3  0.602 -0.053  0.787
     4      5 -0.126 -0.200  0.208
     5      7  0.416 -0.002 -0.275
```

6	8	0.368	0.128	-0.097
7	9	-0.380	-0.259	-0.060
8	10	-0.116	-0.215	-0.204
9	11	0.075	0.701	0.350
10	12	0.170	-0.031	0.113
	etc.			
	etc.			
1261	1756	-0.275	-0.093	0.206
1262	1757	-0.005	-0.028	0.170
1263	1758	-0.166	-0.345	0.055
1264	1759	-0.449	-0.334	-0.234
1265	1760	-0.337	-0.363	-0.291
1266	1761	-0.017	-0.083	-0.151
1267	1762	-0.160	0.353	-0.302
1268	1763	0.285	0.215	0.260
1269	1764	-0.135	-0.328	-0.005
1270	1765	-0.046	-0.161	-0.292

c. BLACK28.DAT

Below is the output for the issue scale parameters -- the **W** matrix and the vector **c** -- for one, two, and three dimensions. After the name of the issue, the first number is the number of respondents placing themselves on the issue, the next number is the constant, **c**, and the last number is r-square. The numbers between **c** and the r-square are the elements of **W**. Note that this is the source for the parameter estimates shown in Table 4B in the **AJPS** article. The slight differences between these numbers and those in the article are due to some marginal improvements made in the program code. The slight differences in the 2nd and 3rd dimensions are due to a slight rotation. The r-squares are unchanged except for a few rounding differences.

LIBERAL/CO	875	4.280	-3.028	0.414		
DEFENSE	1163	5.210	-1.753	0.123		
GOVT SERVI	1119	4.323	4.306	0.451		
INFLATION	816	4.106	2.018	0.159		
ABORTION	1238	2.856	0.624	0.030		
TAX CUTS	836	2.839	-1.074	0.055		
LIBERAL/CO	949	4.369	-2.755	0.414		
GOVT HELP	1160	4.542	-3.400	0.412		
RUSSIA	1152	3.891	-3.034	0.231		
WOMENS EQU	1223	2.845	-2.859	0.204		
GOVT JOBS	1131	4.377	-4.491	0.518		
EQUAL RIGH	1144	2.663	-3.295	0.381		
BUSING	1219	6.051	-2.699	0.256		
ABORTION	1246	2.675	0.722	0.047		
LIBERAL/CO	875	4.300	-2.966	-0.954	0.424	
DEFENSE	1163	5.214	-1.779	-0.899	0.147	
GOVT SERVI	1119	4.368	4.331	-3.042	0.617	
INFLATION	816	4.152	2.088	-2.940	0.393	
ABORTION	1238	2.856	0.512	2.211	0.290	
TAX CUTS	836	2.818	-1.103	0.667	0.071	
LIBERAL/CO	949	4.377	-2.758	-0.459	0.423	
GOVT HELP	1160	4.535	-3.456	0.119	0.424	
RUSSIA	1152	3.887	-3.140	-0.241	0.247	
WOMENS EQU	1223	2.872	-2.466	-6.007	0.771	
GOVT JOBS	1131	4.350	-4.595	2.417	0.635	
EQUAL RIGH	1144	2.673	-3.148	-2.438	0.491	
BUSING	1219	6.049	-2.741	-0.059	0.263	
ABORTION	1246	2.676	0.629	2.112	0.318	
LIBERAL/CO	875	4.294	-2.976	0.708	1.180	0.448
DEFENSE	1163	5.200	-1.806	1.586	-2.562	0.315
GOVT SERVI	1119	4.410	4.295	3.707	-2.929	0.778
INFLATION	816	4.169	1.998	3.286	-1.111	0.451
ABORTION	1238	2.856	0.497	-2.004	-1.174	0.312
TAX CUTS	836	2.813	-1.049	-0.902	0.891	0.091
LIBERAL/CO	949	4.367	-2.785	0.265	0.557	0.437
GOVT HELP	1160	4.534	-3.457	0.140	-0.961	0.440
RUSSIA	1152	3.831	-3.255	1.558	-5.590	0.695
WOMENS EQU	1223	2.891	-2.372	5.602	2.868	0.805
GOVT JOBS	1131	4.341	-4.632	-2.176	-1.392	0.648
EQUAL RIGH	1144	2.680	-3.159	1.860	2.372	0.563
BUSING	1219	6.042	-2.819	0.329	-1.282	0.306
ABORTION	1246	2.675	0.587	-1.980	-0.906	0.329

Appendix B: BMC2.FOR (Bootstrap Program for BLACKBOX.FOR)

1. Introduction

BMC2.FOR is a FORTRAN program that is identical to BLACKBOX.FOR only it performs a bootstrap analysis to estimate the standard errors of the elements of **W** and **c**. The program samples respondents with replacement and re-runs the estimation for each constructed matrix. For example, below 100 bootstrap process is repeated 100 times and the standard errors for **W** and **c** are obtained by computing the sum of squared differences between the actual **W** and **c** and the 100 **W/c**'s from the bootstrap analysis. Because **W** is only defined up to an arbitrary rotation, this was removed before the standard error computation. That is, each **W** from the bootstrap was rotated to best fit the actual **W** before the squared difference was computed.

2. Input File for BMC2.FOR: BMCSTR.DAT

This file is identical to the start file for BLACKBOX.FOR -- BTSTR.DAT -- shown above in Appendix A. The only difference is in the third line where the 5^{th} number is the number of bootstrap iterations.

```
\BLACKB\NES\NES1980.DAT
 DECOMPOSITION OF 14 1980 7-POINT SCALES
    3   14    4    8  100
(8X,I4,527X,I1,13X,I1,11X,I1,11X,I1,11X,I1,13X,I1,1217X,I1,36X,I1,17X,I1,
17X,I1,17X,I1,18X,I1,5X,I1,2X,I1)
LIBERAL/CONSERVATIVE
    0    8    9
DEFENSE
    0    8    9
GOVT SERVICES
    0    8    9
INFLATION
    0    8    9
ABORTION
    0    7    8    9
TAX CUTS
    0    8    9
LIBERAL/CONSERVATIVE
    0    8    9
GOVT HELP MINORITIES
    0    8    9
RUSSIA
    0    8    9
WOMENS EQUAL ROLE
    0    8    9
GOVT JOBS
    0    8    9
EQUAL RIGHTS AMEND
    0    8    9
BUSING
    0    8    9
ABORTION
    0    7    8    9
```

3. Output files for Bootstrap Program

a. BMC223.DAT File

This output file is the same as BLACK23.DAT except that it will contain the output for 300 runs of the program -- 100 one dimensional outputs, 100 two dimensional outputs, and 100 three dimensional outputs. Consequently this file will be rather large and is only useful for debugging purposes if there is some oddity in the data set.

```
Iseed=    93600
 \BLACKB\NES\NES1980.DAT
 DECOMPOSITION OF 14 1980 7-POINT SCALEs
    3   14    4    8  101
(8X,I4,527X,I1,13X,I1,11X,I1,11X,I1,11X,I1,13X,I1,1217X,I1,36X,I1,17X,I1,17X,I1,17X,I1,18X,
I1,5X,I1,2X,I1)
LIBERAL/CO
    0    8    9    0
DEFENSE
    0    8    9    0
GOVT SERVI
    0    8    9    0
INFLATION
    0    8    9    0
ABORTION
    0    7    8    9
TAX CUTS
    0    8    9    0
LIBERAL/CO
    0    8    9    0
GOVT HELP
    0    8    9    0
RUSSIA
    0    8    9    0
WOMENS EQU
    0    8    9    0
GOVT JOBS
    0    8    9    0
EQUAL RIGH
    0    8    9    0
BUSING
    0    8    9    0
ABORTION
    0    7    8    9
 NUMBER OF CASES  1270
 ******************************************************************************
 ******************************************************************************
 ******************************************************************************
 NUMBER OF DIMENSIONS=   1
 ******************************************************************************
 NUMBER OF ROWS                =  1270
 NUMBER OF COLUMNS             =    14
 TOTAL NUMBER OF DATA ENTRIES = 15271
 NUMBER MISSING ENTRIES        =  2509
 PERCENT MISSING DATA          =       14.11136
 SUM OF SQUARES GRAND MEAN     =    52705.13281
 ******************************************************************************
```

Etc., Etc., Etc.

```
 DIMENSION=  1 TOTAL SSE REG1=     25390.1230
 DIMENSION=  1 TOTAL SSE REG2=     25098.9512
 DIMENSION=  1 TOTAL SSE REG1=     25014.0723
 DIMENSION=  1 TOTAL SSE REG2=     24985.0508
 DIMENSION=  1 TOTAL SSE REG1=     24973.4492
 DIMENSION=  1 TOTAL SSE REG2=     24968.3242
 DIMENSION=  1 TOTAL SSE REG1=     24965.9434
 DIMENSION=  1 TOTAL SSE REG2=     24964.7285
 DIMENSION=  2 TOTAL SSE REG1=     20486.9941
 DIMENSION=  2 TOTAL SSE REG2=     20256.3594
```

```
 DIMENSION=  2 TOTAL SSE REG1=     20113.1563
 DIMENSION=  2 TOTAL SSE REG2=     20025.7383
 DIMENSION=  2 TOTAL SSE REG1=     19972.8457
 DIMENSION=  2 TOTAL SSE REG2=     19940.5918
 DIMENSION=  2 TOTAL SSE REG1=     19920.7480
 DIMENSION=  2 TOTAL SSE REG2=     19908.3516
 DIMENSION=  3 TOTAL SSE REG1=     17202.4023
 DIMENSION=  3 TOTAL SSE REG2=     17016.2832
 DIMENSION=  3 TOTAL SSE REG1=     16902.0215
 DIMENSION=  3 TOTAL SSE REG2=     16823.1484
 DIMENSION=  3 TOTAL SSE REG1=     16767.6309
 DIMENSION=  3 TOTAL SSE REG2=     16726.9844
 DIMENSION=  3 TOTAL SSE REG1=     16697.0645
 DIMENSION=  3 TOTAL SSE REG2=     16674.2578
 DIMENSION=  3 TOTAL SSE REG1=     16647.0781
 DIMENSION=  3 TOTAL SSE REG2=     16432.5684
 DIMENSION=  3 TOTAL SSE REG1=     16404.9863
 DIMENSION=  3 TOTAL SSE REG2=     16382.3418
 DIMENSION=  3 TOTAL SSE REG1=     16363.2178
 DIMENSION=  3 TOTAL SSE REG2=     16345.4531
 DIMENSION=  3 TOTAL SSE REG1=     16330.0176
 DIMENSION=  3 TOTAL SSE REG2=     16315.5293
 DIMENSION=  3 TOTAL SSE REG1=     16302.9590
 DIMENSION=  3 TOTAL SSE REG2=     16291.3379
 SINGULAR VALUES OF ESTIMATED MATRICES
 FIRST COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES
 SECOND COLUMN: REPRODUCED MATRIX -- PSI*W + Jc
 THIRD COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES MINUS THE ORIGINAL COLUMN
MEANS
 FOURTH COLUMN:  PSI*W
  1   546.964   545.703   115.836   115.807
  2    97.009    94.903    93.010    92.977
  3    91.542    90.323    76.499    76.470
  4    74.872    73.101    52.415     0.000
  5    47.111     0.000    47.085     0.000
  6    44.479     0.000    43.660     0.000
 ******************************************************************************
 CONSTRAINT CHECKS ON PSI AND W
     SUM OF COLUMNS OF PSI
          0.0000    0.0000    0.0000
     PSI-TRANSPOSE*PSI
     1  115.8070    0.0000    0.0000
     2    0.0000   92.9768    0.0000
     3    0.0000    0.0000   76.4702
     W-TRANSPOSE*W
     1  115.8070    0.0000    0.0000
     2    0.0000   92.9768    0.0000
     3    0.0000    0.0000   76.4701
 R-SQUARE    15234  0.684
 RANK CHECK OF PSI*W    3
    1  115.8070
    2   92.9769
    3   76.4702
    4    0.0000
    5    0.0000
    6    0.0000
 RANK OF SCHOENMANN ROTATION MATRIX    3
 COVARIANCE MATRIX
    1  112.4401    0.9420   -0.3295   -0.0632
    2   76.7452   -0.0914   -0.4333    0.8966
    3   64.1138    0.3228    0.8389    0.4383
 ******************************************************************************
 ITERATION RECORD
   DIM     ERROR    EXPLAINED      PERCENT  CUM PERCENT   R-SQUARE  STD ERR EST
     1  25646.3789  25988.1484     50.3309     50.3309      0.5141      1.3526
     2  20455.4668  31179.0605     10.0532     60.3841      0.6048      1.2736
     3  16291.4453  35343.0820      8.0644     68.4485      0.6845      1.1956
 ******************************************************************************
 ******************************************************************************
```

b. BMC228.DAT File

This output file is the same as BLACK28.DAT except that it will contain the output for 300 runs of the program -- 100 one dimensional outputs, 100 two dimensional outputs, and 100 three dimensional outputs.

```
LIBERAL/CO    875  4.280 -3.028  0.414
DEFENSE      1163  5.210 -1.753  0.123
GOVT SERVI   1119  4.323  4.306  0.451
INFLATION     816  4.106  2.018  0.159
ABORTION     1238  2.856  0.624  0.030
TAX CUTS      836  2.839 -1.074  0.055
LIBERAL/CO    949  4.369 -2.755  0.414
GOVT HELP    1160  4.542 -3.400  0.412
RUSSIA       1152  3.891 -3.034  0.231
WOMENS EQU   1223  2.845 -2.859  0.204
GOVT JOBS    1131  4.377 -4.491  0.518
EQUAL RIGH   1144  2.663 -3.295  0.381
BUSING       1219  6.051 -2.699  0.256
ABORTION     1246  2.675  0.722  0.047
```

Etc.

Etc.

```
LIBERAL/CO    875  4.300 -2.966 -0.954  0.424
DEFENSE      1163  5.214 -1.779 -0.899  0.147
GOVT SERVI   1119  4.368  4.331 -3.042  0.617
INFLATION     816  4.152  2.088 -2.940  0.393
ABORTION     1238  2.856  0.512  2.211  0.290
TAX CUTS      836  2.818 -1.103  0.667  0.071
LIBERAL/CO    949  4.377 -2.758 -0.459  0.423
GOVT HELP    1160  4.535 -3.456  0.119  0.424
RUSSIA       1152  3.887 -3.140 -0.241  0.247
WOMENS EQU   1223  2.872 -2.466 -6.007  0.771
GOVT JOBS    1131  4.350 -4.595  2.417  0.635
EQUAL RIGH   1144  2.673 -3.148 -2.438  0.491
BUSING       1219  6.049 -2.741 -0.059  0.263
ABORTION     1246  2.676  0.629  2.112  0.318
```

Etc.

Etc.

```
LIBERAL/CO    875  4.294 -2.976  0.708  1.180  0.448
DEFENSE      1163  5.200 -1.806  1.586 -2.562  0.315
GOVT SERVI   1119  4.410  4.295  3.707 -2.929  0.778
INFLATION     816  4.169  1.998  3.286 -1.111  0.451
ABORTION     1238  2.856  0.497 -2.004 -1.174  0.312
TAX CUTS      836  2.813 -1.049 -0.902  0.891  0.091
LIBERAL/CO    949  4.367 -2.785  0.265  0.557  0.437
GOVT HELP    1160  4.534 -3.457  0.140 -0.961  0.440
RUSSIA       1152  3.831 -3.255  1.558 -5.590  0.695
WOMENS EQU   1223  2.891 -2.372  5.602  2.868  0.805
GOVT JOBS    1131  4.341 -4.632 -2.176 -1.392  0.648
EQUAL RIGH   1144  2.680 -3.159  1.860  2.372  0.563
BUSING       1219  6.042 -2.819  0.329 -1.282  0.306
ABORTION     1246  2.675  0.587 -1.980 -0.906  0.329
```

b. BMC229.DAT File

This is the important file. This file reports the original **W**'s and **c**'s for one, two, and three dimensions (note that these will be identical to BLACK28.DAT above), the bootstrap standard errors, and the corresponding t-values for the null hypotheses that the parameters are all equal to zero. The r-squares are not shown. As I noted above, the slight differences between these values and those shown in the **AJPS** article for the 2nd and 3rd dimensions are due to a slight rotational difference due to some efficiencies introduced into the program code.

For one dimension, the first number is **c**, the second is **w**, the third and fourth are the standard errors, and the fifth and sixth are the t-values. The output for two and three dimensions is in the same order.

```
  1   4.280  -3.028   0.050   0.143  85.378 -21.137
  2   5.210  -1.753   0.044   0.207 119.542  -8.458
  3   4.323   4.306   0.057   0.191  76.194  22.539
  4   4.106   2.018   0.055   0.223  74.261   9.061
  5   2.856   0.624   0.028   0.134 101.583   4.667
  6   2.839  -1.074   0.047   0.189  60.889  -5.674
  7   4.369  -2.755   0.034   0.138 129.931 -19.985
  8   4.542  -3.400   0.045   0.151 100.279 -22.482
  9   3.891  -3.034   0.054   0.218  72.384 -13.888
 10   2.845  -2.859   0.050   0.271  57.178 -10.558
 11   4.377  -4.491   0.055   0.130  79.867 -34.511
 12   2.663  -3.295   0.043   0.124  61.819 -26.644
 13   6.051  -2.699   0.043   0.222 139.673 -12.131
 14   2.675   0.722   0.024   0.124 110.989   5.807
  1   4.300  -2.966  -0.954   0.050   0.161   0.264  85.190 -18.441
     -3.609
  2   5.214  -1.779  -0.899   0.045   0.213   0.477 116.915  -8.355
     -1.886
  3   4.368   4.331  -3.042   0.052   0.168   0.458  84.058  25.740
     -6.640
  4   4.152   2.088  -2.940   0.052   0.201   0.312  79.912  10.364
     -9.428
  5   2.856   0.512   2.211   0.032   0.124   0.152  90.355   4.127
     14.512
  6   2.818  -1.103   0.667   0.054   0.198   0.393  52.389  -5.579
      1.698
  7   4.377  -2.758  -0.459   0.042   0.144   0.226 104.907 -19.149
     -2.032
  8   4.535  -3.456   0.119   0.049   0.147   0.307  93.011 -23.427
      0.388
  9   3.887  -3.140  -0.241   0.053   0.320   0.915  72.717  -9.802
     -0.263
 10   2.872  -2.466  -6.007   0.049   0.143   0.243  58.883 -17.215
    -24.716
 11   4.350  -4.595   2.417   0.054   0.125   0.303  80.519 -36.869
      7.979
 12   2.673  -3.148  -2.438   0.045   0.141   0.352  58.948 -22.276
     -6.935
 13   6.049  -2.741  -0.059   0.045   0.230   0.380 135.150 -11.931
     -0.156
 14   2.676   0.629   2.112   0.029   0.106   0.130  91.571   5.951
     16.201
  1   4.294  -2.976   0.708   1.180   0.046   0.178   0.509   0.844
     92.487 -16.709   1.391   1.399
  2   5.200  -1.806   1.586  -2.562   0.049   0.441   1.520   2.899
    106.219  -4.099   1.043  -0.884
  3   4.410   4.295   3.707  -2.929   0.064   0.220   1.092   2.305
     68.705  19.511   3.396  -1.270
  4   4.169   1.998   3.286  -1.111   0.052   0.253   0.879   1.495
```

	80.561	7.900	3.737	-0.743				
5	2.856	0.497	-2.004	-1.174	0.028	0.136	0.388	0.633
	101.767	3.642	-5.162	-1.854				
6	2.813	-1.049	-0.902	0.891	0.049	0.316	0.776	1.057
	56.947	-3.321	-1.162	0.842				
7	4.367	-2.785	0.265	0.557	0.043	0.162	0.286	0.611
	102.567	-17.144	0.925	0.912				
8	4.534	-3.457	0.140	-0.961	0.053	0.219	0.867	1.506
	85.959	-15.770	0.161	-0.638				
9	3.831	-3.255	1.558	-5.590	0.098	0.398	0.995	4.111
	38.911	-8.175	1.566	-1.360				
10	2.891	-2.372	5.602	2.868	0.054	0.170	0.541	1.335
	53.952	-13.988	10.348	2.149				
11	4.341	-4.632	-2.176	-1.392	0.053	0.230	0.824	1.201
	81.543	-20.106	-2.641	-1.159				
12	2.680	-3.159	1.860	2.372	0.051	0.145	0.455	1.026
	52.695	-21.720	4.085	2.312				
13	6.042	-2.819	0.329	-1.282	0.046	0.355	1.333	2.688
	132.768	-7.949	0.247	-0.477				
14	2.675	0.587	-1.980	-0.906	0.028	0.128	0.298	0.474
	94.502	4.582	-6.638	-1.913				

Appendix C: BLACKT.FOR

1. Introduction

BLACKT.FOR is a FORTRAN program that can be used to do an Aldrich-McKelvey scaling in more than one dimension. The program reads a "control card" file (BTSTRT.DAT) and the data file (usually an NES data set) and writes three output files: BLACKT23.DAT, BLACKT24.DAT, and BLACKT28.DAT. The program has been compiled for both the Pentium P5 processor as well as the Pentium P6 (Pentium II) processor. These executables are BLACKT5.EXE and BLACKT6.EXE respectively. They will run under both Windows 95 and Windows NT.

The example below uses the liberal-conservative 7-point scale from the 1980 NES cross-sectional survey data set -- NES1980.DAT.

2. Input File for the Transpose Program: BTSTRT.DAT

The first line of the input file (see next page) gives the name of the data set being analyzed. In this case, the 1980 NES data is in the subdirectory \NES. That is, the program, BLACKT6.EXE is in the root directory. Note that if NES1980.DAT could be placed in the same directory as the program the first line would simply be NES1980.DAT.

The second line of the input file is the title of the scaling

The third line contains, in order, the number of basic dimensions, the number of stimuli being placed on the 7-point scale, and the number of missing data values. Note that this is fixed format, namely, in FORTRAN syntax, 3I5. Hence, when you analyze a different scale be sure to ***not change the spacing***. For example, if you want 3 basic dimensions and there are 10 stimuli and 3 missing data values, this line would be:

```
    3   10    3
```

The fourth line contains the values of the missing data. In this case 0, 8, and 9. Note that these numbers are also fixed format.

The fifth line is the format statement for the data file. The 4I1 and the 2I1 are the six stimuli and the I4 is the respondent ID number. This format statement can be figured out by using the ICPSR codebook for the election study. If you have problems figuring out how to do this, just send me E-Mail at kp2a@andrew.cmu.edu.

Finally, the last group of lines are the names of the stimuli. In this case, the four names of the important presidential candidates in 1980 along with the two political parties.

Note that this starting file differs from those for BLACKBOX.FOR and BMC2.FOR in that this program only analyzes *a single issue scale*. Hence, there is no need to repeat the missing data entries.

```
\NES\NES1980.DAT
 DECOMPOSITION OF 1980 LIBERAL-CONSERVATIVE 7-POINT SCALE
    2    6    3
    0    8    9
(8X,I4,1810X,4I1,11X,2I1)
 CARTER
 REAGAN
 KENNEDY
 ANDERSON
 REPUB
 DEMO
```

3. Output files for Transpose X_0 Example

a. BLACKT23.DAT File

This file is the primary output file for the transpose program. For ease of exposition I will annotate this file for the convenience of the reader. My comments will be preceded by #### signs.

The first part of the output file just echoes the lines in BTSTRT.DAT. This is convenient because you can glance at this to make sure the starting file is configured correctly.

```
\BLACKB\NES\NES1980.DAT
 DECOMPOSITION OF 1980 LIBERAL-CONSERVATIVE 7-POINT SCALE
    2    6    3
    0    8    9
(8X,I4,1810X,4I1,11X,2I1)
 CARTER
 REAGAN
 KENNEDY
 ANDERSON
 REPUB
 DEMO

#### Here 888 respondents have been included in the analysis.  The
#### program requires that a respondent place at least s+2 (where s is
#### the number of basic dimensions being estimated) stimuli on the
#### scale.

 NUMBER OF CASES   888
 ***********************************************************************
 ***********************************************************************
 ***********************************************************************

#### The program first estimates a one dimensional model, then two
#### dimensions, etc.

 NUMBER OF DIMENSIONS=   1
 ***********************************************************************

#### This information is only given once.
#### The matix is 6 by 888 and contains 4973 entries.  Hence, there
#### are 355 missing entries or [355/(6*888)]*100 = 6.66291% missing
#### entries.  The Sum of Squares is computed around the grand mean of
#### the matrix.  Hence, it is the sum of the squared differences between
#### the 4973 non-missing entries and the matrix mean.

 NUMBER OF ROWS              =     6
 NUMBER OF COLUMNS           =   888
 TOTAL NUMBER OF DATA ENTRIES =  4973
 NUMBER MISSING ENTRIES      =   355
 PERCENT MISSING DATA        =        6.66291
 SUM OF SQUARES              =    13967.70215
 ***********************************************************************
```

This is the iteration record for the first basic dimension.
REG1 estimates W and c and REG2 estimates ψ.

```
 DIMENSION=  1 TOTAL SSE REG1=      3438.4109
 DIMENSION=  1 TOTAL SSE REG2=      3435.7090
 DIMENSION=  1 TOTAL SSE REG1=      3435.3950
 DIMENSION=  1 TOTAL SSE REG2=      3435.3579
 DIMENSION=  1 TOTAL SSE REG1=      3435.3455
 DIMENSION=  1 TOTAL SSE REG2=      3435.3394
 DIMENSION=  1 TOTAL SSE REG1=      3435.3452
 DIMENSION=  1 TOTAL SSE REG2=      3435.3481
 DIMENSION=  1 TOTAL SSE REG1=      3435.3479
 DIMENSION=  1 TOTAL SSE REG2=      3435.3486
```

The singular values for the one dimensional estimation are reported
below. Only the first **s+3** singular values are shown.

```
 SINGULAR VALUES OF ESTIMATED MATRICES
 FIRST COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES
 SECOND COLUMN: REPRODUCED MATRIX -- PSI*W +  Jc
 THIRD COLUMN:  PSI*W
  1   301.190   301.097    95.444
  2    65.736    64.998     0.000
  3    36.408     0.000     0.000
  4    31.730     0.000     0.000
 **********************************************************************
```

Below are the constraint checks discussed in the AJPS article --
namely, the sum of the columns of $\boldsymbol{\psi}$ must equal zero, and:
$\boldsymbol{\psi}'\boldsymbol{\psi}$ = **W'W** = $\boldsymbol{\Lambda}$ where $\boldsymbol{\Lambda}$ is the s by s diagonal matrix of the
singular values of $\boldsymbol{\psi}$**W'** which is the least squares estimate of
$[\mathbf{X}_0 .- \mathbf{J}_p \underline{\mathbf{c}}']$.

```
 CONSTRAINT CHECKS ON PSI AND W
     SUM OF COLUMNS OF PSI
          0.0000
     PSI-TRANSPOSE*PSI
     1   95.4437
     W-TRANSPOSE*W
     1   95.4438
```

Here $\boldsymbol{\psi}\mathbf{W}' + \mathbf{J}_p\underline{\mathbf{c}}'$ is constructed and the r-square between the elements
of $\boldsymbol{\psi}\mathbf{W}' + \mathbf{J}_p\underline{\mathbf{c}}'$ and the original data matrix, of $\mathbf{X}_0$ is computed as a
check on the estimation.
####

```
 R-SQUARE CHECK    4973  0.754
```

Similar to the above, as a check, a singular value decomposition
of the estimated matrix, $\boldsymbol{\psi}$**W'**, is performed. The rank of the
matrix is reported (here it is one) along with the singular values.

```
 RANK CHECK OF PSI*W    1
    1   95.4437
    2    0.0000
    3    0.0000
    4    0.0000
    5    0.0000
```

```
    6    0.0000
 **********************************************************************
 NUMBER OF DIMENSIONS=   2
 **********************************************************************

#### This is the iteration record for two basic dimensions.  The first
#### basic dimension is extracted first followed by the second basic
#### dimension.

 DIMENSION=  1 TOTAL SSE REG1=      3438.4109
 DIMENSION=  1 TOTAL SSE REG2=      3435.7090
 DIMENSION=  1 TOTAL SSE REG1=      3435.3950
 DIMENSION=  1 TOTAL SSE REG2=      3435.3579
 DIMENSION=  1 TOTAL SSE REG1=      3435.3455
 DIMENSION=  1 TOTAL SSE REG2=      3435.3394
 DIMENSION=  1 TOTAL SSE REG1=      3435.3452
 DIMENSION=  1 TOTAL SSE REG2=      3435.3481
 DIMENSION=  2 TOTAL SSE REG1=      1915.0280
 DIMENSION=  2 TOTAL SSE REG2=      1912.3396
 DIMENSION=  2 TOTAL SSE REG1=      1910.8296
 DIMENSION=  2 TOTAL SSE REG2=      1909.8190
 DIMENSION=  2 TOTAL SSE REG1=      1909.0571
 DIMENSION=  2 TOTAL SSE REG2=      1908.4723
 DIMENSION=  2 TOTAL SSE REG1=      1908.0084
 DIMENSION=  2 TOTAL SSE REG2=      1907.6342
 DIMENSION=  2 TOTAL SSE REG1=      1894.3529
 DIMENSION=  2 TOTAL SSE REG2=      1893.7955
 DIMENSION=  2 TOTAL SSE REG1=      1893.4442
 DIMENSION=  2 TOTAL SSE REG2=      1893.1843
 DIMENSION=  2 TOTAL SSE REG1=      1892.9827
 DIMENSION=  2 TOTAL SSE REG2=      1892.8132
 DIMENSION=  2 TOTAL SSE REG1=      1892.6793
 DIMENSION=  2 TOTAL SSE REG2=      1892.5618
 DIMENSION=  2 TOTAL SSE REG1=      1892.4615
 DIMENSION=  2 TOTAL SSE REG2=      1892.3750

#### The singular values of the estimated matrices are printed out
#### below.  Note that if there were no missing data, the first  s+1
#### singular values in the first two columns would be identical.

 SINGULAR VALUES OF ESTIMATED MATRICES
 FIRST COLUMN:  ORIGINAL MATRIX WITH FILLED IN MISSING ENTRIES
 SECOND COLUMN: REPRODUCED MATRIX -- PSI*W + Jc
 THIRD COLUMN:  PSI*W
  1   301.159   301.131    95.383
  2    65.403    65.256    61.191
  3    60.451    60.393     0.000
  4    30.425     0.000     0.000
  5    24.299     0.000     0.000
 **********************************************************************
 CONSTRAINT CHECKS ON PSI AND W
     SUM OF COLUMNS OF PSI
          0.0000    0.0000
     PSI-TRANSPOSE*PSI
     1   95.3833    0.0000
     2    0.0000   61.1914
     W-TRANSPOSE*W
     1   95.3833    0.0000
     2    0.0000   61.1914
```

```
R-SQUARE CHECK    4973  0.865
RANK CHECK OF PSI*W    2
   1   95.3833
   2   61.1914
   3    0.0000
   4    0.0000
   5    0.0000
   6    0.0000
**********************************************************************
ITERATION RECORD
  DIM     ERROR     EXPLAINED      PERCENT  CUM PERCENT   R-SQUARE
    1   3435.3513  10532.3506     75.4050     75.4050      0.7541
    2   1892.3855  12075.3164     11.0467     86.4517      0.8645
**********************************************************************
```

b. BLACKT24.DAT

Rather than writing out **Y**, for purposes of comparison with the Aldrich-McKelvely procedure, the singular vectors are written out instead. Namely, $\mathbf{Y} = \mathbf{U}\mathbf{L}^{1/2}$, where **U** is p by s matrix such that $\mathbf{U'U} = \mathbf{I_s}$, and **L** is the s by s diagonal matrix of singular values of **Y W'**. Here, just **U** is written out for one and two dimensions.

```
 CARTER      0.242
REAGAN     -0.580
KENNEDY     0.478
ANDERSON    0.059
REPUB      -0.519
DEMO        0.321
CARTER      0.229  0.409
REAGAN     -0.582  0.099
KENNEDY     0.482 -0.001
ANDERSON    0.077 -0.864
REPUB      -0.521  0.097
DEMO        0.315  0.259
```

c. BLACKT28.DAT

This file contains the respondent linear transformation parameters -- **W** and **c**. The first number is the respondent's identification number from the NES1980.DAT file. The second number is the number of responses, the third number is **c**, and the last number is the r-square. Between **c** and the r-square are the **W** values.

```
     1    6  4.333 -0.317  0.351
     8    6  5.167 -0.207  0.463
     9    6  3.833 -0.371  0.884
    10    6  4.167 -0.059  0.048
    11    6  5.167  0.121  0.084
    13    6  5.000 -0.321  0.613
    14    6  4.167 -0.375  0.903
    16    6  4.333 -0.407  0.679
    17    4  3.651  0.204  0.364
    19    6  4.333 -0.310  0.686
    20    6  3.333 -0.407  0.913
                   etc
                   etc
  1756    6  3.667 -0.573  0.888
  1757    6  3.500 -0.115  0.134
  1758    6  3.500 -0.442  0.866
  1759    6  3.500 -0.533  0.722
  1760    5  3.363 -0.322  0.879
  1761    6  4.333 -0.104  0.140
  1762    5  2.586 -0.229  0.283
  1764    6  3.833 -0.387  0.962
  1765    5  4.020  0.169  0.340
     1    6  4.333 -0.311 -0.364  0.636
     8    6  5.167 -0.202 -0.281  0.989
     9    6  3.833 -0.371  0.041  0.893
    10    6  4.167 -0.064  0.229  0.521
    11    6  5.167  0.119  0.093  0.111
    13    6  5.000 -0.323  0.059  0.635
    14    6  4.167 -0.372 -0.145  0.979
    16    6  4.333 -0.410  0.209  0.805
    17    4  3.845  0.202  0.283  0.985
    19    6  4.333 -0.306 -0.180  0.817
    20    6  3.333 -0.408  0.011  0.914
                  etc
                  etc
  1756    6  3.667 -0.569 -0.252  0.984
  1757    6  3.500 -0.110 -0.282  0.630
  1758    6  3.500 -0.445  0.186  0.978
  1759    6  3.500 -0.525 -0.427  0.999
  1760    5  2.827 -0.349  0.386  0.969
  1761    6  4.333 -0.108  0.222  0.559
  1762    5  2.597 -0.219 -0.380  0.948
  1764    6  3.833 -0.388  0.075  0.992
  1765    5  3.173  0.127  0.628  0.636
```

Footnotes

[1] See note 4 of **AJPS** article.

[2] See note 11 of **AJPS** article.