DOCUMENT RESUME

ED 455 278 TM 033 081

AUTHOR	Fan, Xitao
TITLE	Using Commonly Available Software for Conducting Bootstrap Analyses.
PUB DATE	2001-04-00
NOTE	41p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001).
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Computer Software; Educational Research; Heuristics; *Nonparametric Statistics; Psychological Studies; Reliability; *Research Methodology; Statistical Inference
IDENTIFIERS	*Bootstrap Methods

ABSTRACT

Bootstrap analysis, both for nonparametric statistical inference and for describing sample results stability and replicability, has been gaining prominence among quantitative researchers in educational and psychological research. Procedurally, however, it is often quite a challenge for quantitative researchers to implement bootstrap analysis in their research because bootstrap analysis is typically not an automated program option in statistical software programs. This paper uses a few heuristic analytical examples to show how bootstrap analysis can be accomplished through the use of some commonly available statistical software programs. Until bootstrap analysis becomes an automated program option in standard statistical software programs (e.g., the Statistical Package for the Social Sciences or the Statistical Analysis System), quantitative researchers may have to make do with these or other creative approaches to accomplish bootstrap analysis in their research. (Contains 4 tables, 10 figures, and 37 references.) (Author/SLD)

Using Commonly Available Software for Conducting Bootstrap Analyses

Xitao Fan

University of Virginia

Running Head: Conducting Bootstrapping

Send correspondence about this paper to:

Xitao Fan
Curry School of Education
University of Virginia
405 Emmet Street South
PO Box 400277
Charlottesville, VA 22904-4277
Phone: (804)243-8906
Fax: (804)924-1384
E-Mail: xfan@virginia.edu
Web: http://www.people.virginia.edu/~xf8d

Paper presented at the 2001 annual meeting of the American Educational Research Association, April 10-14 (Session #: 39.01), Seattle, Washington.

Abstract

Bootstrap analysis, both for non-parametric statistical inference, and for describing sample results stability and replicability, has been gaining prominence among quantitative researchers in educational and psychological research. Procedurally, however, it is often quite a challenge for quantitative researchers to implement bootstrap analysis in their research, because bootstrap analysis is typically not an automated program option in statistical software programs. This paper uses a few heuristic analytical examples to show how bootstrap analysis can be accomplished through the use of some commonly available statistical software programs. Until bootstrap analysis becomes an automated program option in standard statistical software programs (e.g., SPSS, SAS), quantitative researchers may have to make do with these or other creative approaches to accomplish bootstrap analysis in their research.

KEY WORDS: bootstrap, resampling, computing-intensive, non-parametric inference, sample results replicability

In educational and psychological research and measurement, the traditional over-reliance on statistical significance testing has been challenged on several grounds. These include issues related to sample size, the meaningfulness of the traditional null hypothesis, and questions involving the validity of theoretical assumptions underlying parametric statistical inferences (e.g., Carver, 1978; Shaver, 1993; Thompson, 1993). As a result of these and other concerns about traditional statistical significance tests, researchers are increasingly turning to empirically-grounded resampling procedures.

Bootstrap is probably the best-known resampling method, and it has been applauded by some as one of the newest breakthroughs in statistics (Kots & Johnson, 1992). The significance of bootstrapping as a versatile analytic approach to data analysis has been widely recognized not only by those in the area of statistics, but also by quantitative researchers in social and behavioral sciences. In the area of education, the recognition for bootstrap method was evidenced by the invited keynote address delivered by the pioneer in bootstrapping, Bradley Efron, at the 1995 AERA Annual Meeting (Efron, 1995).

Instead of relying upon the <u>theoretical assumptions</u> to derive sampling distributions for statistical estimators, the bootstrap method estimates these distributions empirically, using information drawn from the sample of observations used to estimate the statistical model in the first place (Diaconis & Efron, 1983; Efron, 1979). In doing so, the bootstrap avoids some of the pitfalls of traditional statistical significance testing. As discussed by Lunnenborg (2000), "Until inexpensive computing power made replicate data analysis practical, the drawing of statistical inferences from a set of data almost always required that we accept an idealized model for the origin of those data. Such models can be either inappropriate or inadequate for the data in our study. Resampling techniques allow us to base the analysis of a study solely on the design of that study, rather than on a poorly-fitting model" (p. xi).

In social and behavioral sciences, bootstrap method has been used in a variety of research situations, and for many different statistical techniques. For example, bootstrap method has been applied in psychological measurement for such issues as differential test predictive validity (e.g., Fan & Mathews, 1994) and item bias (e.g., Harris & Kolen, 1989), and in sociological research (e.g., Stine, 1989). The application of bootstrapping method has involved many different statistical techniques, including correlation analysis (e.g., Mendoza, Hart, & Powell, 1991; Rasmussen, 1987), regression analysis (e.g., Fan & Jacoby, 1995), discriminant analysis (e.g., Dalgleish, 1994), canonical correlation analysis (e.g., Fan & Wang, 1996), factor analysis (e.g., Lambert, Wildt, & Durand, 1991; Thompson, 1988), and structural equation modeling (e.g., Bollen & Stine, 1990; Yung & Bentler, 1996).

In addition to using bootstrap for nonparametric statistical inference (Efron, 1985), bootstrap method has also been advocated as a descriptive tool and an internal replication approach for assessing the stability and replicability of sample results of an individual study (Thompson, 1993). This descriptive use of bootstrap is meaningful when our interest may not be about statistical inference, but rather, about getting an understanding about how stable/unstable the sample results may be.

Bootstrapping is a computing-intensive data resampling strategy, and widespread access to powerful computing facilities makes bootstrap estimation an attractive and viable procedure for research practitioners. Unfortunately, although the logic of bootstrapping is conceptually straightforward to understand, bootstrapping in substantive research has yet to enjoy widespread use. Because bootstrapping is not typically implemented in the major commercial statistical software packages such as SAS and SPSS, researchers who desire to use this strategy will have to write their own programs for performing bootstrap resampling. This can be a daunting endeavor for researchers who do not have the programming skills, knowledge, or interest required to carry out such a task. Consequently, this has become a major obstacle for implementing bootstrapping in substantive research,

Some methodologists have sensed the need for programs that can be used by other research practitioners to perform bootstrapping; as a result, some special programs have been published for different analytic techniques, such as regression analysis (Fan & Jacoby, 1995), and factor analysis (Thompson, 1988). But overall, bootstrapping remains procedurally difficult for most research practitioners, because the commonly used major statistical software packages (e.g., SAS, SPSS) have not implemented bootstrapping as a program analysis option. Many research practitioners are not aware, however, that bootstrap analysis has been implemented in some widely available, although more specialized, software programs, and a little creativity is all that is needed for taking advantage of these program features. Furthermore, bootstrap analysis can also be accomplished by using a standard statistical analysis package (e.g., SAS) with only a reasonable amount of effort. This paper attempts to provide some heuristic examples of implementing bootstrap analysis for some common statistical techniques by using some widely available statistical software programs.

The paper has two sections. The first section provides bootstrapping examples within the framework of structural equation modeling. For quantitative researchers who use some of the widely available structural equation modeling programs (e.g., AMOS, EQS. LISREL), it will be shown that many statistical techniques can be implemented as a structural equation model, and subsequently, bootstrap analysis can be accomplished by using the bootstrap routine already built in these structural equation modeling programs.

The second section provides some bootstrap analysis examples in SAS. For quantitative researchers who use SAS system for statistical analysis, this section will become especially handy for implementing bootstrap analysis for a variety of statistical techniques with only a relatively small amount of effort. For the purpose of illustration, several heuristic statistical analysis examples will be used in this paper, and a heuristic data set is created and used for these analyses.

Bootstrap Through Structural Equation Modeling Programs

Many different statistical techniques exist, and these different techniques may be designed for different purposes, and may have different historical origins. But as many quantitative researchers are aware, many commonly used statistical procedures are variations of the general linear model (GLM) (Knapp, 19xx). In the past few decades, structural equation modeling (SEM) has increasingly been seen as a useful quantitative technique for specifying, estimating, and testing hypothesized models describing relationships among a set of substantively meaningful variables. Much of SEM's attractiveness is due to its applicability in a wide variety of research situations, a versatility that has been amply demonstrated (e.g., Bollen & Long, 1993; Byrne, 1994; Jöreskog & Sörbom, 1989; Loehlin, 1992).

Many widely used statistical techniques may be considered as special cases of SEM, including regression analysis, canonical correlation analysis, confirmatory factor analysis, and path analysis (Bagozzi, Fornell & Larcker, 1981; Bentler, 1992; Fan, 1996; Jöreskog & Sörbom, 1989). Because of such generality, SEM has been considered as a unified model which joins methods from econometrics, psychometrics, sociometrics, and multivariate statistics (Bentler, 1994). In short, for researchers in the social and behavioral sciences, SEM has become an important tool for testing theories with both experimental and non-experimental data (Bentler & Dudgeon, 1996). Within the general framework of GLM, some seemingly different statistical techniques can be implemented as a structural equation model, and the analysis can be accomplished using any of the available structural equation modeling programs (e.g., AMOS, EQS, LISREL, PROC CALIS in SAS).

Despite the fact that bootstrapping has not been implemented in any of the standard statistical software packages (e.g., SAS, SPSS), it has been available from some of the software packages designed for structural equation modeling for some time. Because of the versatility of SEM as a general analytic approach for statistical techniques in the GLM family, practically, many analytical techniques within the GLM family can be translated to SEM models, and thus conducted by using SEM approach. As a result, bootstrap analysis can be accomplished for some commonly used statistical techniques through some of these more specialized SEM software packages.

Several structural equation modeling software programs have implemented bootstrap routine as an option, and EQS is probably a pioneer to do so (Bentler, 1992)[1]. More recently, AMOS has almost perfected its bootstrap routine by including a fully automated bootstrap analysis option, including different bootstrap options, and automated bootstrap analysis results output, thus making bootstrap analysis very accessible and user-friendly (Arbuckle & Wothke, 1999). Because this paper is not about structural equation modeling per se, but rather, it is about implementing bootstrapping by using SEM software, easiness in implementing bootstrapping is of the major concern. As a result, I will primarily rely on AMOS for illustrating how bootstrap

[1] This is when I first acquired EQS software, and bootstrap was already implemented in EQS. This was the earliest among structural equation modeling software programs I was aware of. EQS may have implemented bootstrap earlier than that indicated here, but I have not investigated earlier versions of EQS about this.

analysis can be accomplished through structural equation modeling approach. An EQS program example for conducting bootstrap analysis will also be provided.

A Heuristic Data Set

Table 1 presents a heuristic data set that will be used in bootstrap analysis examples in this paper. This heuristic data set contains a continuous variable Y, three other continuous variables X_1, X_2, and X_3, and a categorical variable G with three levels (A, B, C). The three levels of the categorical variable G are also represented by two dummy coded variables D_1 and D_2. The heuristic data set has very small sample size (n=24). This data set will serve as the "parent" sample from which bootstrap samples will be obtained. Ideally, the size of the "parent" sample in bootstrap analysis should be larger, so that the sample is more likely to be representative, and that it is less likely to obtain duplicate bootstrap samples (i.e., identical bootstrapped samples). For the simplicity of our illustration, however, we will ignore these and other similar issues associated with small size of the "parent" sample.

Insert Table 1 about here

Example 1: Correlation Analysis

Correlation analysis is very common among educational researchers. Within the framework of structural equation modeling, applying bootstrapping for correlation analysis is straightforward. Assuming that for the three variables X_1, X_2, and X_3 in the heuristic data set in Table 1, we are interested in the correlations among the three variables. We are further interested in conducting bootstrap analysis for the sample correlation coefficients.

Using the data provided in Table 1, we can obtain the sample correlations among the three variables (X_1, X_2, and X_3) as follows:

X_1	1.0		
X_2	.7632	1.0	
X_3	.4928	.6829	1.0

In structural equation modeling, the correlation analysis for the three observed variables is essentially a saturated confirmatory factor analysis model with three correlated factors, each with its own sole error-free indicator. Because it is saturated model, the model has perfect fit. Using AMOS Graphics, a window graphic interface for model specification, the correlation model can be specified in AMOS as shown in Figure 1. The heuristic data set in Table 1 can be read by AMOS in a variety of formats, such as ASCII data, and SPSS system file. One simply needs to open the data set and specify the data format from the "DATA FILES" under "FILE" of AMOS Graphics.

Insert Figure 1 about here

Once the model is specified as in Figure 1, and the data set opened by AMOS, one needs to request bootstrap routine for the specified model analysis. This is easily accomplished by selecting "ANALYSIS PROPERTIES" under "VIEW/SET' of AMOS, as shown in Figure 2.

Insert Figure 2 about here

Once "ANALYSIS PROPERTIES" is selected, a variety of analysis options can be specified. One tab under "ANALYSIS PROPERTIES" is "Bootstrap", and click this tab will show the options available for performing bootstrapping analysis for the specified model, as shown in Figure 3. For performing bootstrap analysis, the option "Performing Bootstrap" needs to be checked. One may choose the percentile confidence interval or the bias-corrected confidence interval, and the desired confidence interval level (default is 90%). The details about these bootstrapping confidence intervals are provided elsewhere (e.g., Lunneborg, 2000; Mooney & Duval, 1993), and are not discussed here. Once these options and the estimation method are specified (in Figure 3 example, 200 bootstrap samples[2], 90% percentile confidence intervals, and maximum likelihood estimation) for the bootstrap analysis as shown in Figure 3, we are ready to conduct bootstrap analysis for the model of the three correlations.

Insert Figure 3 about here

To conduct model analysis in AMOS, from AMOS' "MODEL-FIT", select "Calculate Estimates", as shown in Figure 4. This will start the model and bootstrap analysis as specified by the graphic model in Figure 1, and all the options requested for bootstrap analysis previously.

Insert Figure 4 about here

Once the analysis is completed, the regular and bootstrap analysis out put can be viewed under AMOS "View/Set" either in text format output, or in table format output. The following are the selected bootstrap analysis results (highlighted) for the three correlations among X_1, X_2, and X_3 in the heuristic data set in Table 1:

[2] Bootstrap analysis may require more bootstrap samples, e.g., 1000, to achieve estimation stability for the distribution of a statistic of interest. In our heuristic example here, we have a very small "parent" sample, and will only have 200 bootstrapped samples.

Correlations:

	Estimate	SE	Mean	Lower	Upper
x2 <--> x1	0.76	**0.09**	**0.75**	**0.60**	**0.89**
x2 <--> x3	0.68	**0.13**	**0.68**	**0.45**	**0.86**
x1 <--> x3	0.49	**0.19**	**0.50**	**0.17**	**0.78**

It is noticed that, based on 200 bootstrapped samples, the empirical standard error for the correlation coefficient r_{13} (.19) is twice as large as that for r_{12} (.09). Consequently, the 90% confidence interval for r_{13} is much wider (.17, .78) than that for r_{12} (.60, .89).

Example 2: Regression Analysis

Regression analysis is probably one of the most widely utilized statistical technique in educational research. Regression analysis can easily be translated into a structural model, as shown by many (e.g., Bentler, 1992; Jöreskog & Sörbom, 1989). For the heuristic data set, when Y is used as the dependent variable, and X_1, X_2, and X_3 are used as the three predictors, we have the following selected regression analysis results:

REGRESSION: Y1 = X1 X2 X3

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	71.83672	2.26746	31.68	<.0001	0
X1	1	0.35949	0.23032	1.56	0.1343	0.16361
X2	1	0.62252	0.28686	2.17	0.0422	0.27097
X3	1	0.46164	0.06741	6.85	<.0001	0.63503

R-Square 0.9086 F = 66.25, p < .0001

It is shown that, based on statistical theory, and assuming that the data satisfies all the assumptions of regression analysis, X1 does not have statistically significant unique contribution to Y when X2 and X3 are already in the regression model.

Figure 5 presents the AMOS structural equation model representation for the regression analysis conducted above, with three predictors (X_1, X_2, and X_3), and Y as the outcome variable. Again, this is a saturated model, and model-data fit is perfect.

Insert Figure 5 about here

Once the model specification is complete in AMOS, as shown in Figure 5, specification of bootstrap options is done in the same fashion as described in the previous correlation analysis (see Figures 2 and 3, and related discussion). In this analysis, because we are also interested in the replicability of the sample R^2 of the regression model, under the "OUTPUT" tab of AMOS "ANALYSIS PROPERTIES", the option of "SQUARED MULTIPLE CORRELATIONS" was checked, as shown in Figure 6.

Insert Figure 6 about here

Based on 200 bootstrapped samples, the following selected bootstrap analysis results were obtained from AMOS (empirical standard error, bootstrapped mean, 90% confidence interval):

Regression Weights:

	Estimate	S.E.	C.R.	P	SE	Mean	Lower	Upper
y1 <-- x1	0.36	0.21	1.67	0.09	**0.25**	**0.35**	**-.11**	**0.70**
y1 <-- x2	0.62	0.27	2.33	0.02	**0.28**	**0.61**	**0.19**	**1.11**
y1 <-- x3	0.46	0.06	7.34	0.00	**0.07**	**0.47**	**0.38**	**0.61**

Standardized Regression Weights:

	Estimate	SE	Mean	Lower	Upper
y1 <-- x1	0.16	0.12	**0.17**	**-0.05**	**0.34**
y1 <-- x2	0.27	0.11	**0.27**	**0.08**	**0.46**
y1 <-- x3	0.64	0.07	**0.65**	**0.53**	**0.76**

Squared Multiple Correlation ($\underline{R}^2$):

	Estimate	SE	Mean	Lower	Upper
y1	0.91	0.02	**0.93**	**0.89**	**0.96**

For regression coefficients, compared with the standard regression analysis results presented previously, the bootstrapped results are quite similar. For example, the regression coefficient for X_1 could be zero in both analyses, and the theoretically- and empirically-based standard errors (under S.E. and SE respectively) are reasonably comparable. The bootstrap analysis also shows that the model R^2 is very high (bootstrapped mean of $R^2 = .93$), and it is likely to be highly replicable across comparable samples (bootstrapped 90% confidence limits: .89 to .96).

Example 3: Analysis of Variance (ANOVA)

As widely discussed elsewhere, ANOVA is part of the general linear model, and ANOVA analysis can be viewed as a special case of regression analysis. The research question, "Do the three groups, as represented by the three levels of the variable G, have the same population means?", can be readily translated into a regression model through the use of one of the coding scheme for the group membership (i.e., dummy coding, effect coding, or orthogonal coding). In this illustrative example, the simple dummy coding (D_1 and D_2) is used for representing the group membership for the variable G. Using standard regression analysis approach for this ANOVA problem, we obtained the following results:

REGRESSION APPROACH TO ANOVA: Y1 = D1 D2
TWO DUMMY VARIABLES FOR THREE GROUPS

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	104.87500	2.02330	51.83	<.0001	0
D1	1	-11.75000	2.86138	-4.11	0.0005	-0.76894
D2	1	-6.87500	2.86138	-2.40	0.0256	-0.4499
R-Square	0.4477	F = 8.51, p = .002				

The results above show that it is unlikely that the three groups have equal populations means (F=8.51, p=.0020). The group membership accounts for about 45% of the variance in the outcome variable Y (R^2=.4477). The intercept (104.875) represents the mean of Group C (the group coded as 0s on both dummy variables). Furthermore, the coefficients associated with the two dummy coded variables (D_1 and D_2) represent the group mean difference between Groups A and C, and that between Groups B and C, respectively. The statistical tests for the two regression coefficients of the dummy variables represent the tests for the contrasts[3] between the group means of A and C, and that between B and C. These two tests (t=-4.11, p=.0005; t=-2.40, p=.0256) indicate that both Group A and Group B have statistically lower means than Group C.

As shown previously, a regression model is easily translated to a saturated structural equation model. Figure 7 shows the AMOS model for the ANOVA analysis conducted above. By specifying requested output (see Figure 6) and bootstrap options (see Figure 3) in AMOS, we obtain the following selected bootstrap analysis results:

[3] These contrasts are equivalent to Dunnett's test in ANOVA. For details, see Pedhazur (1997, Chapter 11).

Regression Weights:

	Estimate	S.E.	C.R.	P	SE	Mean	Lower	Upper	P
y1 <-- D1	-11.75	2.73	-4.30	0.00	**2.10**	**-11.41**	**-14.93**	**-8.10**	**0.01**
y1 <-- D2	-6.88	2.73	-2.51	0.01	**2.66**	**-6.67**	**-10.84**	**-1.79**	**0.05**

Squared Multiple Correlations:

	Estimate	SE	Mean	Lower	Upper
y1	0.45	**0.13**	**0.49**	**0.28**	**0.72**

Based on 200 bootstrap samples, the bootstrap empirical results related to testing for the group differences (as represented by the regression coefficients, and statistical tests associated with the coefficients) indicate that the groups A (represented by D_1, with 90% confidence limits being –14.93 and –8.10) and B (represented by D_2, with 90% confidence limits being –10.84 and –1.79) have statistically different means from that of Group C. The bootstrapped $\underline{R}^2$ appears to be somewhat larger (mean of $\underline{R}^2 = .49$) than the original sample $\underline{R}^2$ (.45). The bootstrap 90% confidence limits for $\underline{R}^2$ are .28 and .72.

<u>Example 4: Measurement Reliability Analysis</u>

The widely used Cronbach's coefficient $\underline{\alpha}$ provides accurate reliability estimate for composite score that consists of τ-equivalent measures (i.e., measures of the same latent dimension in the same measurement unit, but with possibly different precision). When the composite score consists of more realistic congeneric measures (i.e., measures of the same latent dimensions, but in possibly different measurement units and with possibly different precision), Cronbach's coefficient $\underline{\alpha}$ provides the lower-bound estimate for the composite score reliability. It is probably not widely known that structural equation modeling approach can be used for measurement reliability analysis. As demonstrated Raykov (1997, 1998), such an approach is reasonably straightforward. In addition, the structural equation modeling approach for measurement reliability analysis also has the advantage of providing a more accurate reliability estimate for a composite score consisting of congeneric measures, rather than only the lower-bound estimate as Cronbach's coefficient $\underline{\alpha}$ does in the same situation. For this reason, the reliability estimate from this structural equation modeling approach tends to be higher than Cronbach's coefficient alpha. The more the measures deviate from τ-equivalent measures, the more obvious is the difference between Cronbach's coefficient alpha and that estimated from this structural equation modeling approach.

Assuming that we have a composite consisting of $\underline{k}$ components, and we are interested in the composite score reliability estimate. As shown by Raykov (1997), the model in Figure 8 represents a structural equation model for estimating the reliability for the composite consisting of congeneric measures (X_1 to X_k). The correlation between F_2 (a phantom variable, representing "observed score") and F_1 (representing "true score") is the reliability index. In the model, this is represented by the two asterisks followed by $\sqrt{\rho_{xx}}$, indicating that it is a reliability index, not reliability coefficient itself. The square of this index is the estimated reliability coefficient for the composite of the k-item scale. For the distinction between reliability index and reliability coefficient, see Crocker and Algina (1986). For more details for this model of measurement reliability analysis, see Raykov (1997, 1998). As noted by Raykov, the reliability index to be estimated in the model in Figure 8 is not an inherent parameter of the model, but is readily obtained as the correlation between the two latent factors (F_2 and F_1).

Insert Figure 8 about here

For our heuristic data set in Table 1, we have three X variables (X_1, X_2, and X_3). Assuming that these are scores on three components of a measurement scale, and we are interested in estimating the measurement reliability for the composite consisting of these three components. Using a standard statistical package (e.g., SPSS, SAS), the Cronbach's coefficient alpha for the component consisting of X_1, X_2, and X_3 can be obtained: $\alpha = .647$.

Implementing the structural equation model for measurement reliability could easily lead us to conducting bootstrap analysis for the reliability estimate of the composite consisting of X_1, X_2, and X_3. Figure 9 is the AMOS model for obtaining measurement reliability estimate in the form of reliability index. As discussed above, the measurement reliability index is not an inherent model parameter, but is readily obtained as the correlation between two latent factors of F1 (analogous to "true score") and F2 (a phantom variable analogous to "observed score").

Insert Figure 9 about here

In order to obtain the correlation between the two latent factors as the reliability index, appropriate output option must be selected. Under "OUTPUT" tab in AMOS "ANALYSIS PROPERTIES", the option of "All Implied Moments" is selected for obtaining the correlation between F1 and F2, which is the reliability index for the composite consisting of X_1, X_2, and X_3. Because the correlation between F1 and F2 is sought, the option of "Standardized Estimates" (i.e., standardized covariance) is also selected, as is shown in Figure 10.

Insert Figure 10 about here

Implementing the model in Figure 9 in AMOS, and the bootstrap options we have selected, the reliability index (i.e., the correlation between F1 and F2) is estimated to be .847.

Based on 200 bootstrap samples, the bootstrapped 90% confidence interval has lower limit of .722, and upper limit of .960. If we square these estimates (reliability coefficient is the squared reliability index, see discussion above), we have the estimated reliability coefficient of .717, with bootstrapped 90% lower and upper confidence limits of .521 and .922. The width of the bootstrapped confidence limits (.521 - .922) suggests that there tends to be a lack of stability in sample measurement reliability for the composite score consisting of X_1, X_2, and X_3, and the an estimate based on a single sample is unlikely to be stable in replication.

Compared with Cronbach's $\underline{\alpha}$ of .647, the estimated reliability coefficient of .717 through structural equation modeling approach is higher. As discussed previously, this is expected, because Cronbach's $\underline{\alpha}$ is only the lower bound reliability for a composite that does consist of τ-equivalent measures. For the three X variables used in this example, they are obviously on very different measurement scales, a violation of τ-equivalent measures. As discussed by Raykov (1997), in this situation, the reliability coefficient from the structural equation modeling approach will be higher than the classic Cronbach's $\underline{\alpha}$, and is typically a more accurate estimate of measurement reliability than the classic $\underline{\alpha}$.

Conducting Bootstrapping Analysis in Other Structural Equation Modeling Programs

The examples discussed above have relied on AMOS for implementing the bootstrap analyses, because AMOS is the only program that has automated the bootstrap analysis in terms of specifying bootstrap options and obtaining bootstrap analysis results. With a reasonable amount of effort, the same analyses can be conducted through some other structural equation modeling programs, such as EQS and LISREL. A major difference between using AMOS and other structural equation modeling programs is in the degree of automation. In AMOS, bootstrap analysis is fully automated. In EQS, bootstrap is a built-in program option. If this analysis is requested, the bootstrap analysis results (e.g., parameter estimates) will be exported to an external data set in ASCII format. Subsequently, any standard statistical software (e.g., SPSS, SAS) can be used to access this external ASCII data set and to conduct further descriptive analyses about the bootstrapped estimates (e.g., to obtain the bootstrap confidence interval for a parameter). As an example, an EQS program for conducting bootstrap analysis for multiple regression ($Y = a + b_1{*}X_1 + b_2{*}X_2 + b_3{*}X_3$) is provided in Table 2.

Insert Table 2 about here

Bootstrapping in SAS

SAS is a very flexible system that provides all kinds of capabilities for data management, statistical analysis, statistical programming, etc. Although SAS does not provide any automated program option for conducting bootstrap analysis, bootstrapping can be accomplished through SAS with a reasonable amount of effort, thanks to its built-in flexibility in programming. For quantitative researchers who use SAS as the primary data analysis tool, to get to know how to conduct bootstrap analysis in SAS can become very convenient.

In using SAS for bootstrap analysis, four programming components are needed: 1) draw bootstrap samples from a given "parent" data set through sampling with replacement, 2) for each bootstrapped sample, conduct relevant statistical analysis of interest, 3) extract the needed parameter estimates of interest, accumulate these estimates from all the bootstrapped samples, and store them in a data file, and 4) conduct analysis for the parameter estimates of all bootstrapped samples.

Assuming that we construct a composite measure that consists of X_1, X_2, and X_3 in the heuristic data in Table 1 as its components. We are interested in conducting bootstrap analysis for the measurement reliability estimate in the form of Cronbach's coefficient α for this composite. Table X presents an annotated SAS macro program for bootstrapping Cronbach's coefficient α for the composite consisting of X_1, X_2, and X_3. In this SAS macro program, the purposes of the different clusters of SAS commands are reasonably clear from the comments in the program.

Insert Table 3 about here

It should be noted that for any statistical analysis conducted in SAS, upon request, SAS will output a SAS data set that contains the sample statistics for the analysis conducted. For conducting bootstrap analysis, it is necessary to extract the sample statistic(s) of interest from this SAS data set. In the SAS program in Table 3, the command "OUTP=STATSOUT" requests such an output SAS data set that contains sample Cronbach's coefficient $\underline{\alpha}$, and name this SAS data set as "STATSOUT" for future use.

Because different analytic procedures have different statistics, the SAS data output for sample statistics have different data structure for different statistical analytic procedures . One needs to understand the data structure of this SAS output data set for the statistical analysis implemented. To view the data structure of the SAS output data set, one simply needs to use SAS "PROC PRINT" procedure to display the SAS output data set. To illustrate how the sample statistic(s) of interest can be extracted, the following is the SAS output data set structure from the coefficient alpha analysis under "PROC CORR" in SAS:

Obs	_TYPE_	_NAME_	X1	X2	X3
1	MEAN		10.5417	14.2083	30.7500
2	STD		3.3490	3.2030	10.1221
3	N		24.0000	24.0000	24.0000
4	RAWALPHA		0.6473	0.6473	0.6473
5	STDALPHA		0.8457	0.8457	0.8457
6	RAWALDEL		0.5641	0.4543	0.8652
7	STDALDEL		0.8116	0.6603	0.8657
8	RAWCTDEL		0.5932	0.7807	0.6238
9	STDCTDEL		0.6846	0.8369	0.6261

The shaded entry in the above data set is the sample coefficient alpha that we need to extract. A few straightforward SAS statements in Table 3 SAS program accomplish this task. Finally, the bootstrap sample coefficient alpha is appended to a file of a permanent SAS data set. With each bootstrap iteration, bootstrap sample statistic(s) of interest are accumulated in this permanent SAS data set for future analyses.

It should be noted that adapting the SAS program in Table 3 for other statistical analysis procedures is reasonably straightforward. As a matter of fact, only the shaded SAS statements in the SAS program in Table 3 need to be replaced. Table 4 presents three groups of SAS codes for conducting bootstrap analysis for the heuristic data set in Table 1 for i) correlation analysis among X_1, X_2, and X_3, ii) regression analysis, iii) one-way ANOVA through the use of dummy coded variables. Using each group of SAS codes in Table 4 to replace the shaded SAS statements in Table 3 will provide the bootstrap analysis for the examples discussed previously in this paper as AMOS examples.

Insert Table 4 about here

Conclusions

Bootstrap analysis, both as a tool for non-parametric statistical inference, and as a tool for describing sample results stability and replicability, has been gaining prominence among quantitative researchers in educational and psychological research. Procedurally, it is often quite a challenge for many quantitative researchers to implement bootstrap analysis in their research, because bootstrap analysis is typically not an automated program option in statistical software programs. Creative approaches, however, can be taken to implement bootstrap analysis through some commonly available software programs. This paper shows a few examples of how this can be accomplished. Until bootstrap analysis becomes a standard program option in statistical analysis software programs (e.g., SPSS, SAS), quantitative researchers may have to make do with the approaches illustrated here if bootstrap analysis is desired.

References

Arbuckle, J. L., & Wothke, W. (1999). Amos 4.0 User's Guide. Chicago, IL: Smallwaters Corporation.

Bagozzi, R. P., Fornell, C., & Larcker, D. F. (1981). Canonical correlation analysis as a special case of structural relations model. Multivariate Behavioral Research, 16, 437-454.

Bentler, P. M. (1992). EQS structural equations program manual. Los Angeles, CA: BMDP Statistical Software.

Bentler, P. M. (1994). Forward. In B. M. Byrne, Structural Equation Modeling with EQS and EQS/Windows. Thousand Oaks, CA: SAGE.

Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. Annual Review of Psychology, 47, 563-592.

Bollen, K. A. & Long, J. S. (1993). Introduction. In K. A. Bollen and J. S. Long (Eds.), Testing structural equation models (pp. 1-9). Newbury Park, California: SAGE Publications.

Bollen, K. A., & Stine, R. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In K. A. Bollen & J. S. Long, (Eds.), Testing structural equation models (pp. 111-135). Newbury Park, CA: Sage.

Byrne, B. M. (1994). Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming. Newbury, California: SAGE Publications, Inc.

Carver, R.P. (1978). The case against statistical significance testing. Harvard Educational Review, 48, 378-399.

Crocker, L., & Algina, J. (1986). Introduction to classical & modern test theory. Fort Worth, TX: Holt, Rinehart and Winston, Inc.

Dalgleish, L. I. (1994). Discriminant analysis: Statistical inference using the jackknife and bootstrap procedures. Psychological Bulletin, 116, 498-508.

Diaconis, P., & Efron, B. (1983). Computer-intensive methods in statistics. Scientific American, May, 116-130.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. the Annals of Statistics, 7, 1-26.

Efron, B. (1985). Bootstrap confidence intervals for a class of parametric problems. Biometrika, 72, 45-48.

Efron, B. (April, 1995). Computers, bootstraps, and statistics. Invited address to American Educational Research Association Annual Meeting, Division D--Measurement and Research Methodology. San Francisco, CA.

Fan, X. (1996). Structural equation modeling and canonical correlation analysis: What do they have in common? Structural Equation Modeling: A Multidisciplinary Journal, 4, 64-78.

Fan, X., & Jacoby, W. R. (1995). BOOTSREG: A SAS matrix language program for bootstrapping linear regression models. Educational and Psychological Measurement, 55, 764-768.

Fan, X., & Mathews, T. A. (1994, April). Using bootstrap procedures to assess the issue of predictive bias in college GPA prediction for ethnic groups. Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document Reproduction Service #: ED 372 117).

Fan, X. & Wang, L. (1996). Comparability of jackknife and bootstrap results: An investigation for a case of canonical analysis. Journal of Experimental Education, 64, 173-189.

Harris, D. J., & Kolen, M. J. (1989). Examining the stability of Angoff's delta item bias statistic using bootstrap. Educational and Psychological Measurement, 49, 81-87.

Jöreskog, K. G. & Sörbom, D. (1989). LISREL 7: A Guide to the Program and applications, 2nd Edition. Chicago, IL: SPSS Inc.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance testing system. Psychological Bulletin, 85, 410-416.

Kotz, S., & Johnson, N. L. (1992). Breakthroughs in statistics: Volumes 1 and 2. New York: Springer-Verlag.

Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1991). Approximating confidence interval for factor loadings. Multivariate Behavioral Research, 26, 421-434.

Loehlin, J. C. (1992). Latent variable models: An introduction to factor, path, and structural analysis. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Lunneborg, C. E. (2000). Data analysis by Resampling: Concepts and applications. Pacific Grove, CA: Duxbury Thompson Learning.

Mendoza, J. L., Hart, D. E., & Powell, A. (1991). A bootstrap confidence interval based on a correlation corrected for range restriction. Multivariate Behavioral Research, 26, 255-269.

Mooney, C. Z., & Duval, R. D. (1993). Bootstrapping: A nonparametric approach to statistical inference. Sage University Series on Quantitative Applications in the Social Sciences, 07-095. Newbury Park: CA: Sage.

Pedhazur, E. J. (1997). Multiple regression in behavioral research: Explanation and prediction (3rd ed.). Fort Worth, TX: Harcourt Brace College Publishers.

Rasmussen, J. L. (1987). Estimating the correlation coefficients: Bootstrap and parametric approaches. Psychological Bulletin, 101, 136-139.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. Applied Psychological Measurement, 21, 173-184.

Raykov, T. (1998). A method of obtaining standard errors and confidence intervals of composite reliability for congeneric items. Applied Psychological Measurement, 22, 369-374.

Shaver, J.P. (1993). What significance testing is, and what it isn't. The Journal of Experimental Education, 61, 293-316.

Stine, R. A. (1989). An introduction to bootstrap methods: Examples and ideas. Sociological Methods and Research, 8, 243-291.

Thompson, B. (1988). Program FACSTRAP: A program that computes bootstrap estimates of factor structure. Educational and Psychological Measurement, 48, 681-686.

Thompson, B. (1993). The use of statistical significance tests in research: Bootstrap and other alternatives. The Journal of Experimental Education, 61, 361-377.

Yung, Y., & Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), Advanced structural equation modeling: Issues and techniques (pp. 195-226). Mahwah, New Jersey: Lawrence Erlbaum.

Table 1 A Heuristic Data Set

Y	X_1	X_2	X_3	G	D_1	D_2
93	9	12	20	A	1	0
88	7	10	15	A	1	0
95	8	12	26	A	1	0
95	10	14	21	A	1	0
95	9	12	25	A	1	0
99	10	18	31	A	1	0
99	8	10	34	A	1	0
81	7	9	16	A	1	0
95	5	14	30	B	0	1
88	10	12	15	B	0	1
99	5	11	42	B	0	1
87	9	9	16	B	0	1
101	13	14	29	B	0	1
102	10	15	32	B	0	1
110	18	20	51	B	0	1
102	10	17	31	B	0	1
106	14	18	39	C	0	0
103	12	17	32	C	0	0
103	16	17	34	C	0	0
103	11	14	35	C	0	0
105	12	15	37	C	0	0
107	16	19	39	C	0	0
106	14	16	39	C	0	0
106	10	16	49	C	0	0

Table 2 An EQS Program Example for Conducting Bootstrap Analysis for Regression

```
/TITLE
   EQS BOOTSTRAP EXAMPLE - REGRESSION ANALYSIS
/SPECIFICATIONS
   CAS=24; VAR=7; ME=ML; MA=RAW;
   DATA='C:\BOOTSTRAP\DATA\ASCII.TXT';
/LABELS
   V1=Y1; V2=X1; V3=X2; V4=X3; V5=G; V6=D1; V7=D2;
/EQUATIONS
  V1 = *V2 + *V3 + *V4 + E1;
/VARIANCES
  V2 TO V4=*; E1=*;
/COVARIANCES
  V2 TO V4=*;
/SIMULATION
  REPLICATIONS=200;
  BOOTSTRAP=24;
/OUTPUT
  DATA='C:\BOOTSTRAP\EQS\BTRPOUT.TXT';
  PARAMETER ESTIMATES;
/END
```

Note: In the EQS program above, commands for bootstrap analysis are in **bold face**. Underlined EQS commands specify the name and location of the file that stores each of the 200 bootstrap sample results in ASCII format. Once this EQS analysis is complete, a standard statistical package (e.g., SAS, SPSS) can be used to read the ASCII data set "BTRPOUT.TXT" and to conduct further analyses for the bootstrapped parameter estimates (e.g., constructing bootstrap percentile method confidence interval for a regression coefficient).

Table 3 A SAS Macro Program for Conducting Bootstrap Analysis for Coefficient Alpha

```
LIBNAME BTRAP 'C:\AERA2001\BTRAP';

DATA RAWDATA;  INFILE 'C:\AERA2001\BTRAP\ASCII.TXT';
  INPUT Y X1-X3 G D1 D2;

%LET BTRAPN=200;    * <DEFINE INTENDED NUMBER OF BOOTSTRAPPED SAMPLES>;
%LET SMPLN=24;      * <EACH BOOTSTRAP SAMPLE N, USUALLY ORIGINAL SAMPLE N>;

%MACRO BTRAP;                * <START OF BOOTSTRAP MACRO 'BTRAP'>;
%DO A=1 %TO &BTRAPN;

DATA BTDATA;                 * <BOOTSTRAP SAMPLING WITH REPLACEMENT>;
  DROP I;
  DO I=1 TO &SMPLN;
    IOBS=INT(RANUNI(0)*N) + 1;
    SET RAWDATA POINT=IOBS NOBS=N;
  OUTPUT;
  END;
STOP;

        * <CONDUCTING STATISTICAL ANALYSIS, OUTPUT SAMPLE STATISTICS>;

PROC CORR NOCORR NOPRINT NOMISS ALPHA OUTP=STATSOUT;
  VAR X1 X2 X3;
  RUN; QUIT;

DATA STATS;            * <EXTRACT SAMPLE STATISTICS OF INTEREST>;
  SET STATSOUT;
  IF _TYPE_='RAWALPHA';
  ALPHA=X1;
  KEEP ALPHA;

     * <APPEND EACH BOOTSTRAP SAMPLE STATISTICS TO A SAS DATA SET>;

PROC APPEND DATA=STATS BASE=BTRAP.STATS FORCE;
%END;
%MEND BTRAP;                 * <END OF BOOTSTRAP MACRO>;
%BTRAP;                      * <RUNNING BOOTSTRAP MACRO>;
RUN;

      * <ANALYZE BOOTSTRAPPED DISTRIBUTIONS OF STATISTICS>;

DATA RESULTS;   SET BTRAP.STATS;
PROC UNIVARIATE;
  RUN; QUIT;
```

Note: For conducting bootstrap analysis for other statistical techniques, only the shaded texts in the above need to be changed.

Table 4 SAS Codes for Conducting Bootstrap Analysis for Three Other Statistical Analysis

i. Correlation Analysis among X_1, X_2, and X_3:

```
PROC CORR NOPRINT OUTP=STATSOUT;
  VAR X1 X2 X3;
RUN;

DATA CORR12;
  SET STATSOUT;
  CORR12=X1;
  IF _NAME_ NE 'X2' THEN DELETE;
  KEEP CORR12;
  RUN;
DATA CORR1323;
  SET STATSOUT;
  CORR13=X1; CORR23=X2;
  IF _NAME_ NE 'X3' THEN DELETE;
  KEEP CORR13 CORR23;
  RUN;
DATA STATS;
  MERGE CORR12 CORR1323;
  RUN;
```

ii. Regression Analysis ($Y = X_1 X_2 X_3$):

```
PROC REG NOPRINT OUTEST=STATSOUT ADJRSQ RSQUARE;
  MODEL Y=X1 X2 X3;
  RUN;

DATA STATS;
  SET STATSOUT;
  KEEP INTERCEPT X1 X2 X3 _RSQ_ _ADJRSQ_;
  RUN;
```

iii. One-Way ANOVA ($Y = D_1 D_2$):

```
PROC REG NOPRINT OUTEST=STATSOUT ADJRSQ RSQUARE;
  MODEL Y=D1 D2;
  RUN;

DATA STATS;
  SET STATSOUT;
  KEEP INTERCEPT D1 D2 _RSQ_ _ADJRSQ_;
  RUN;
```

Figure Captions:

Figure 1: Model Representation for Correlation Analysis in AMOS

Figure 2: Selecting "ANALYSIS PROPERTIES" in AMOS to Request Bootstrap Analysis

Figure 3: Specifying Bootstrap Options

Figure 4: Conducting Model Analysis in AMOS

Figure 5: Regression Analysis Model in AMOS

Figure 6: Selecting Output Options AMOS

Figure 7: AMOS Model for the ANOVA (Two Dummy Variables for Three Groups)

Figure 8: Structural Equation Model for Reliability Index of A Scale of k-Components

Figure 9: AMOS Model for Reliability Index of a Composite with Three Components

Figure 10: AMOS Output Option Selection for Obtaining Reliability Index

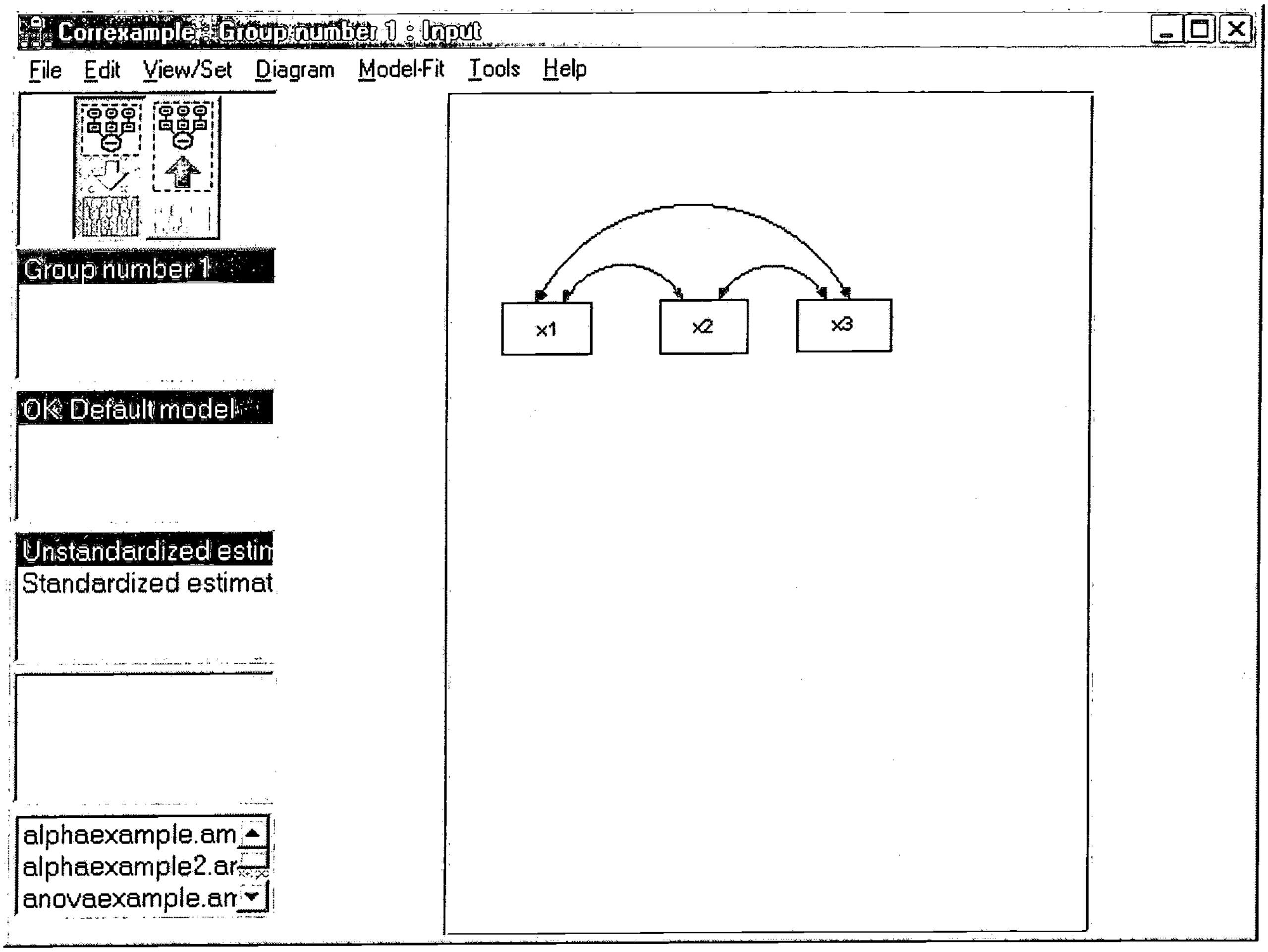

Figure 1: Model Representation for Correlation Analysis in AMOS

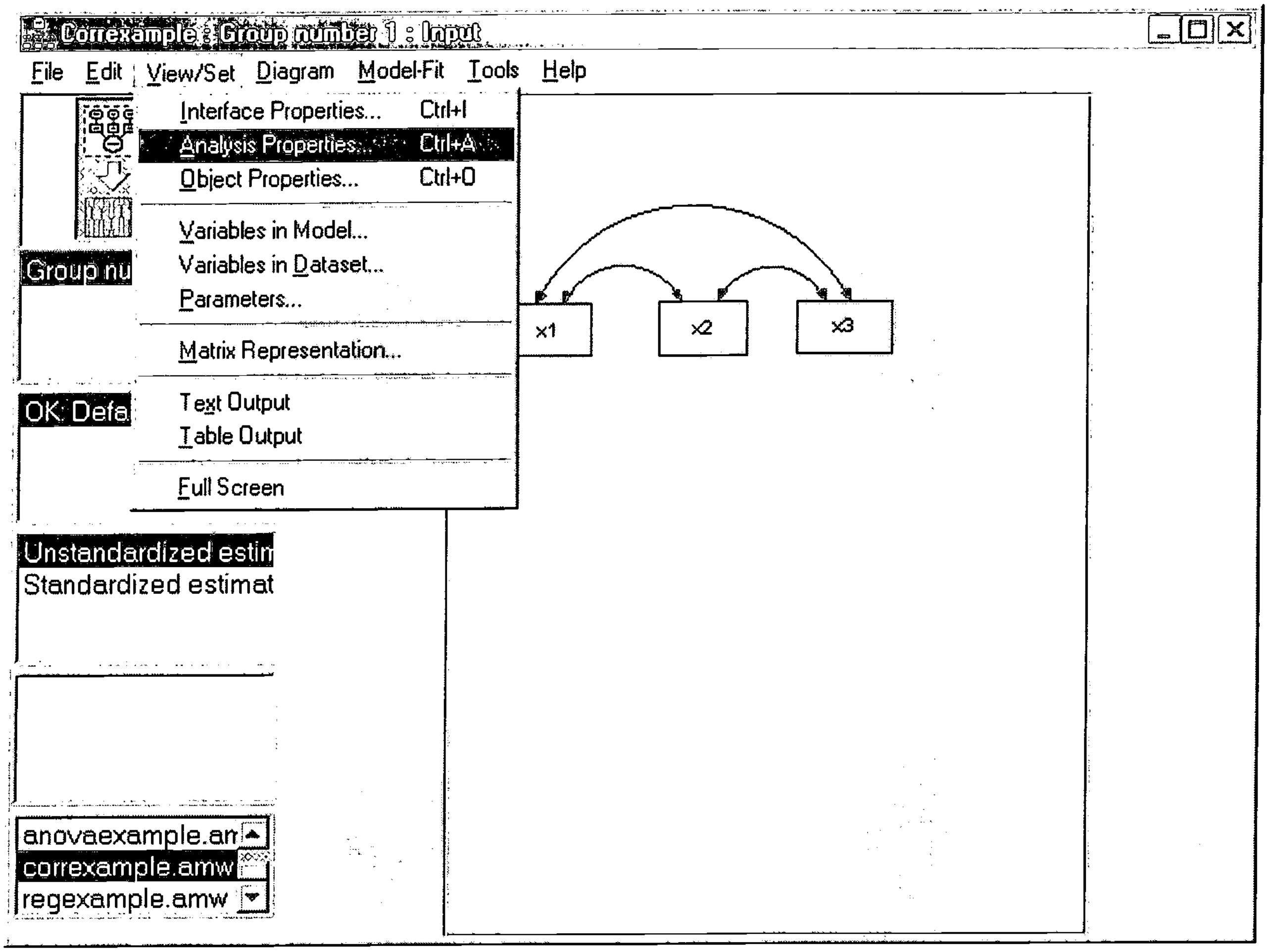

Figure 2 Requesting Bootstrap Analysis Selecting "ANALYSIS PROPERTIES" in AMOS

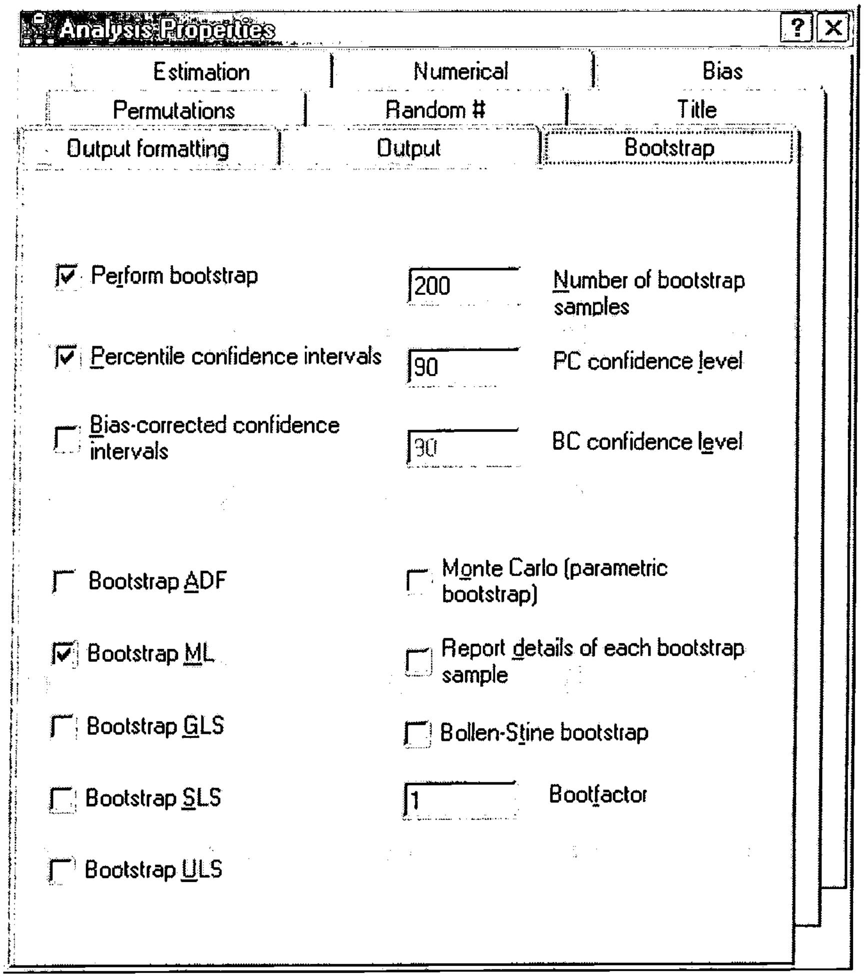

Figure 3 Specifying Bootstrap Options

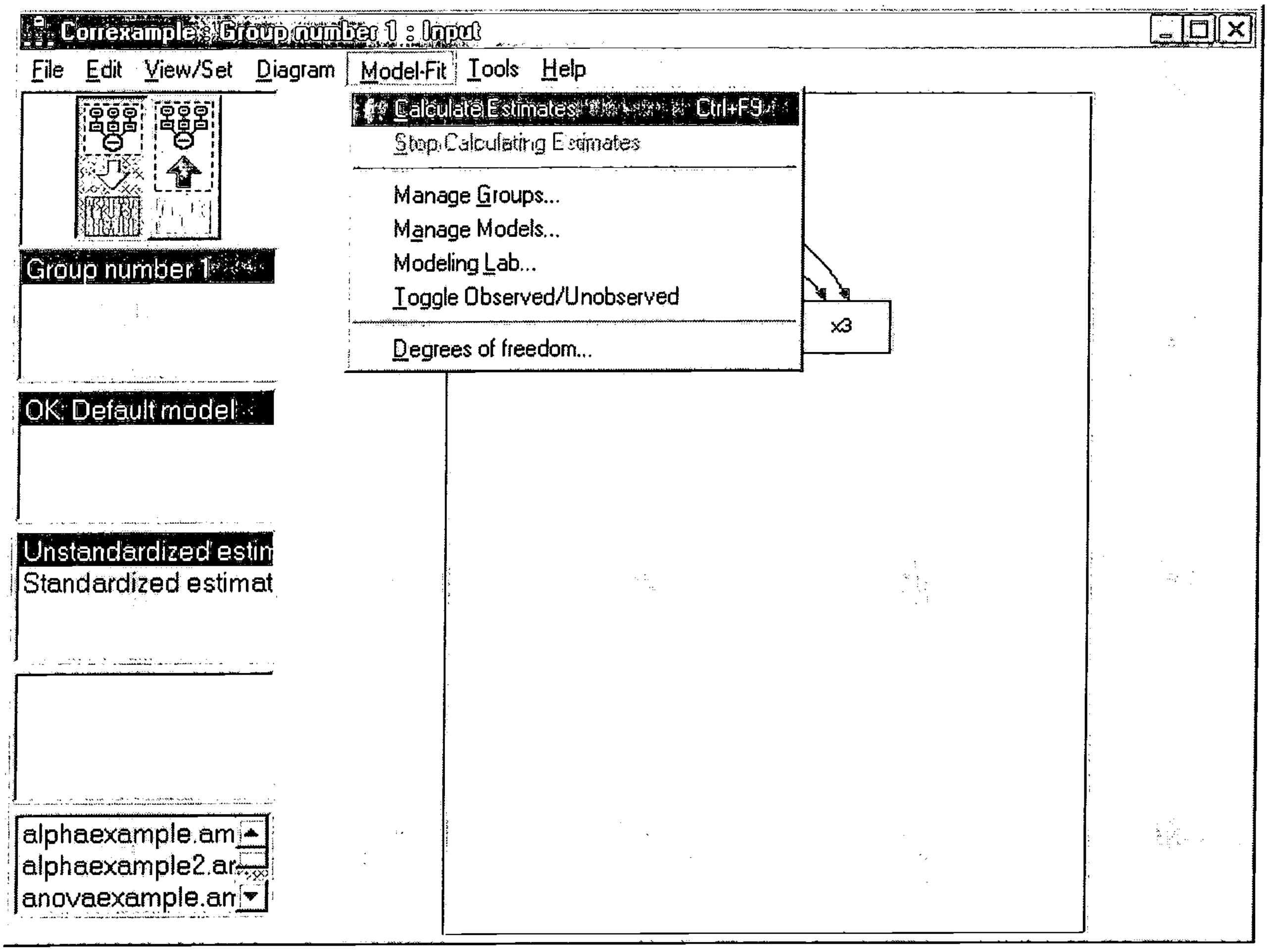

Figure 4 Conducting Model Analysis in AMOS

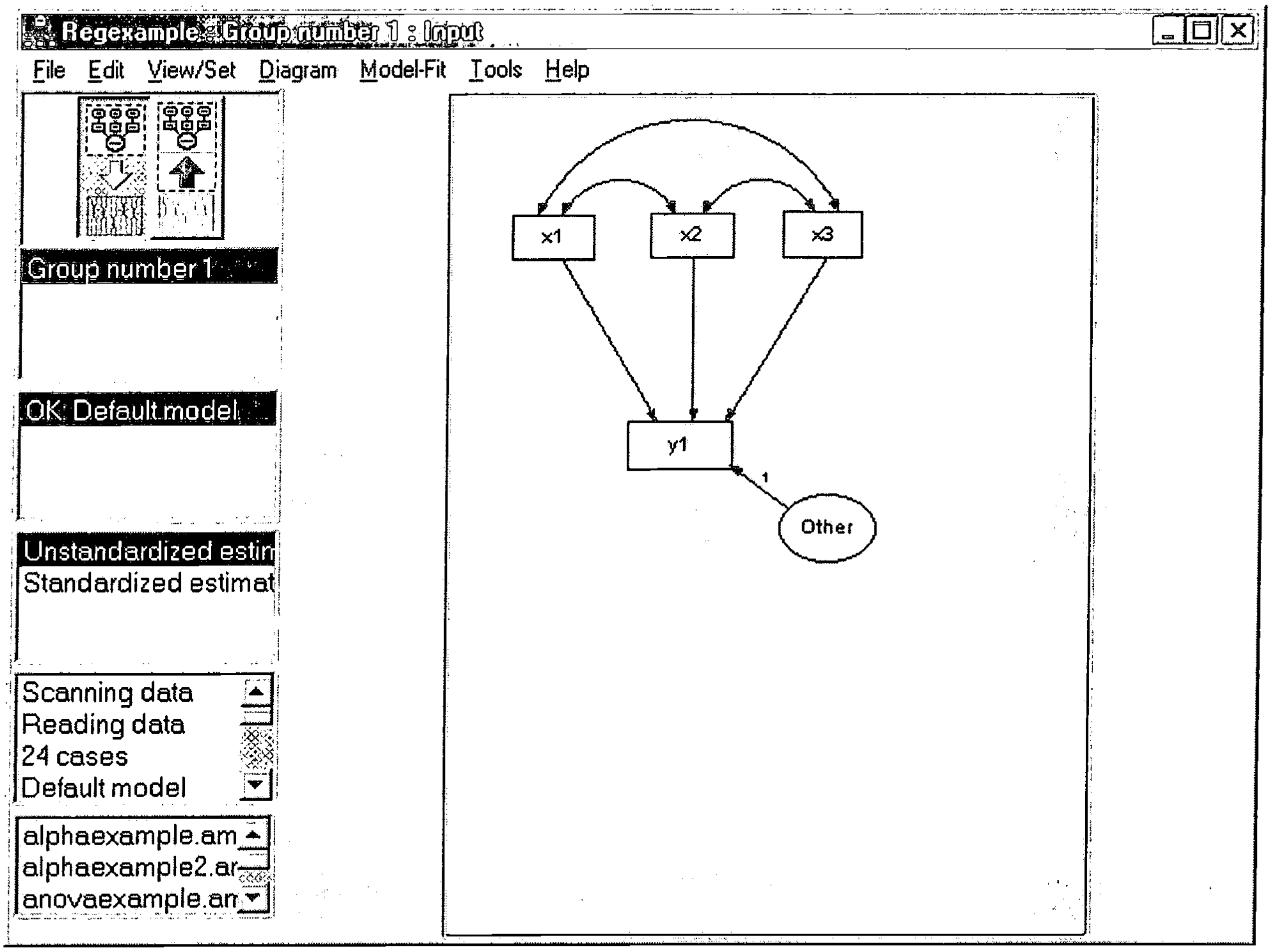

Figure 5 Regression Analysis Model in AMOS

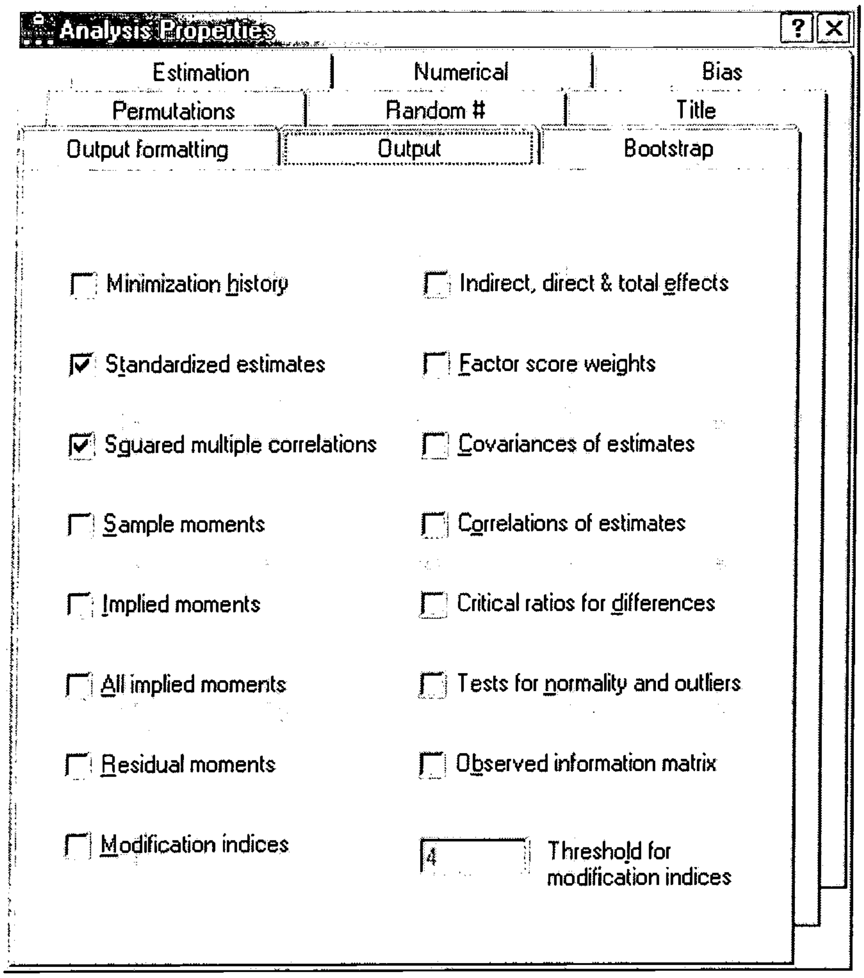

Figure 6 Selecting Output Options AMOS

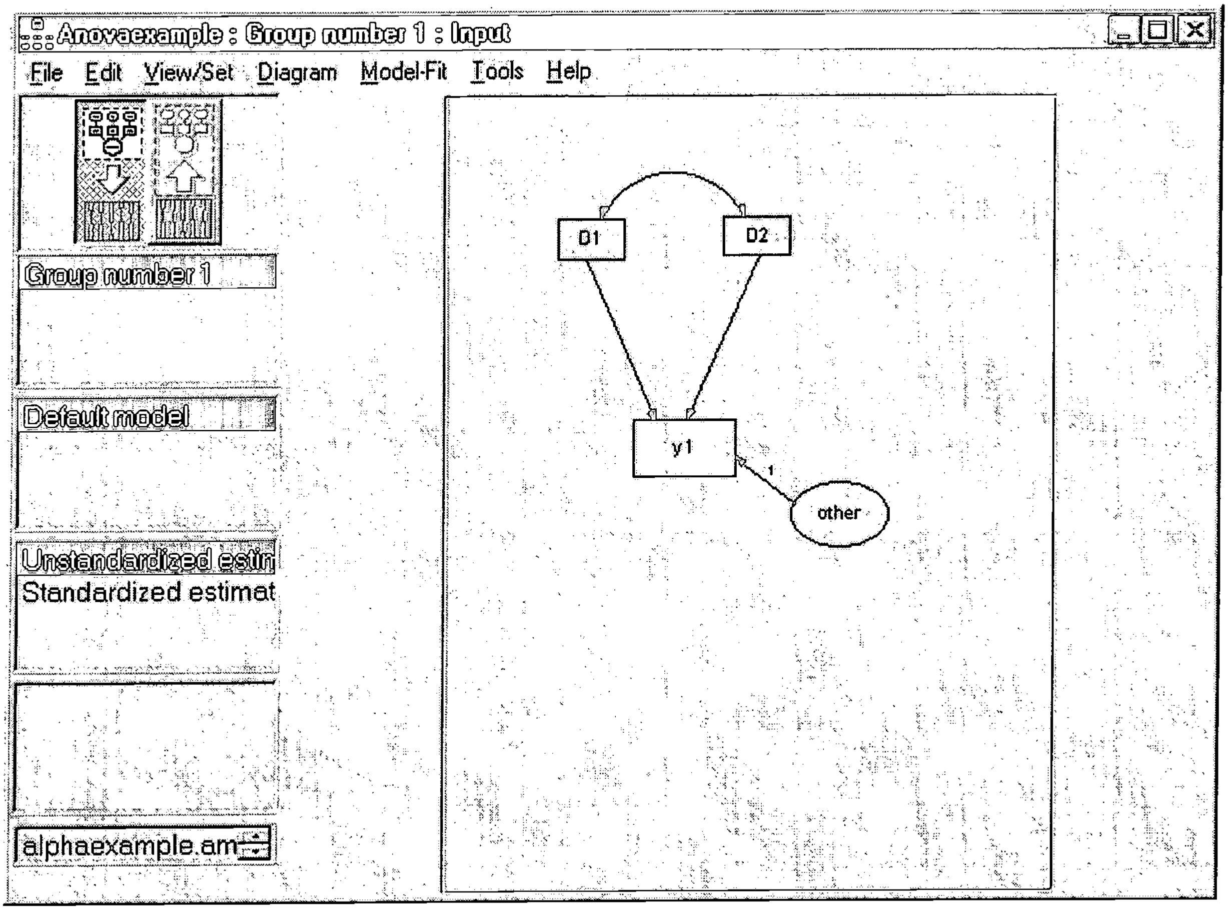

Figure 7 AMOS Model for the ANOVA (Two Dummy Variables for Three Groups)

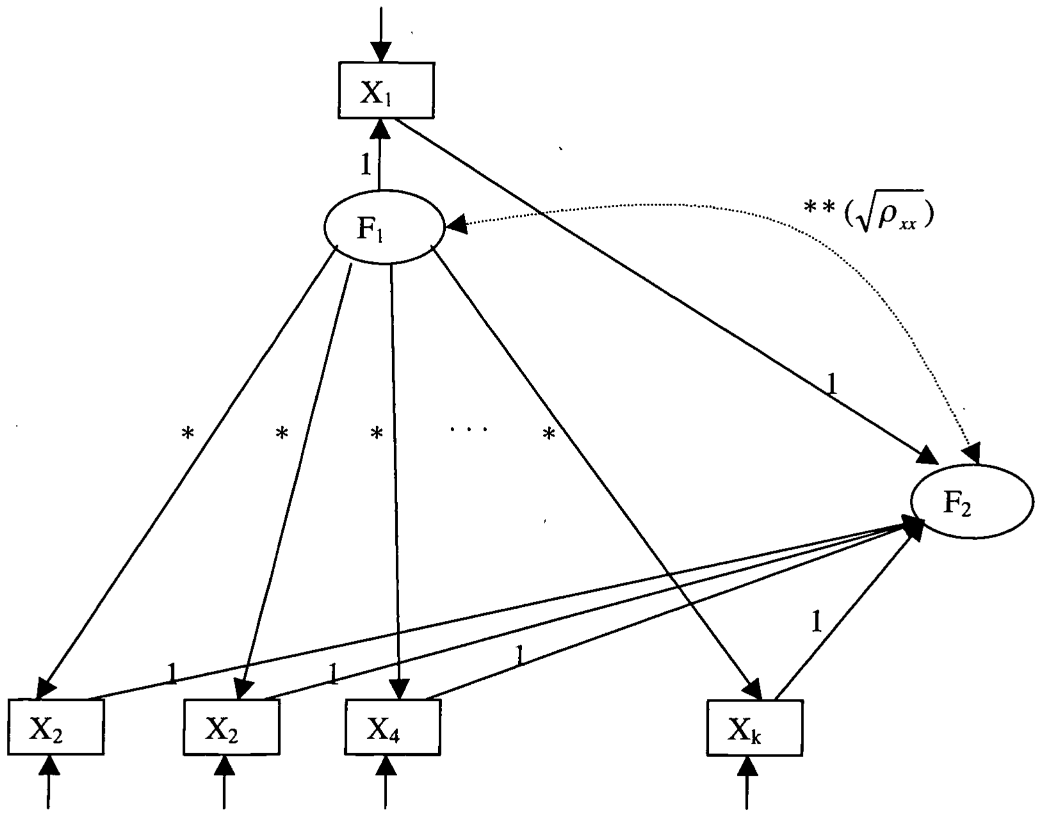

Figure 8 Structural Equation Model for Reliability Index of A Scale of k-Components

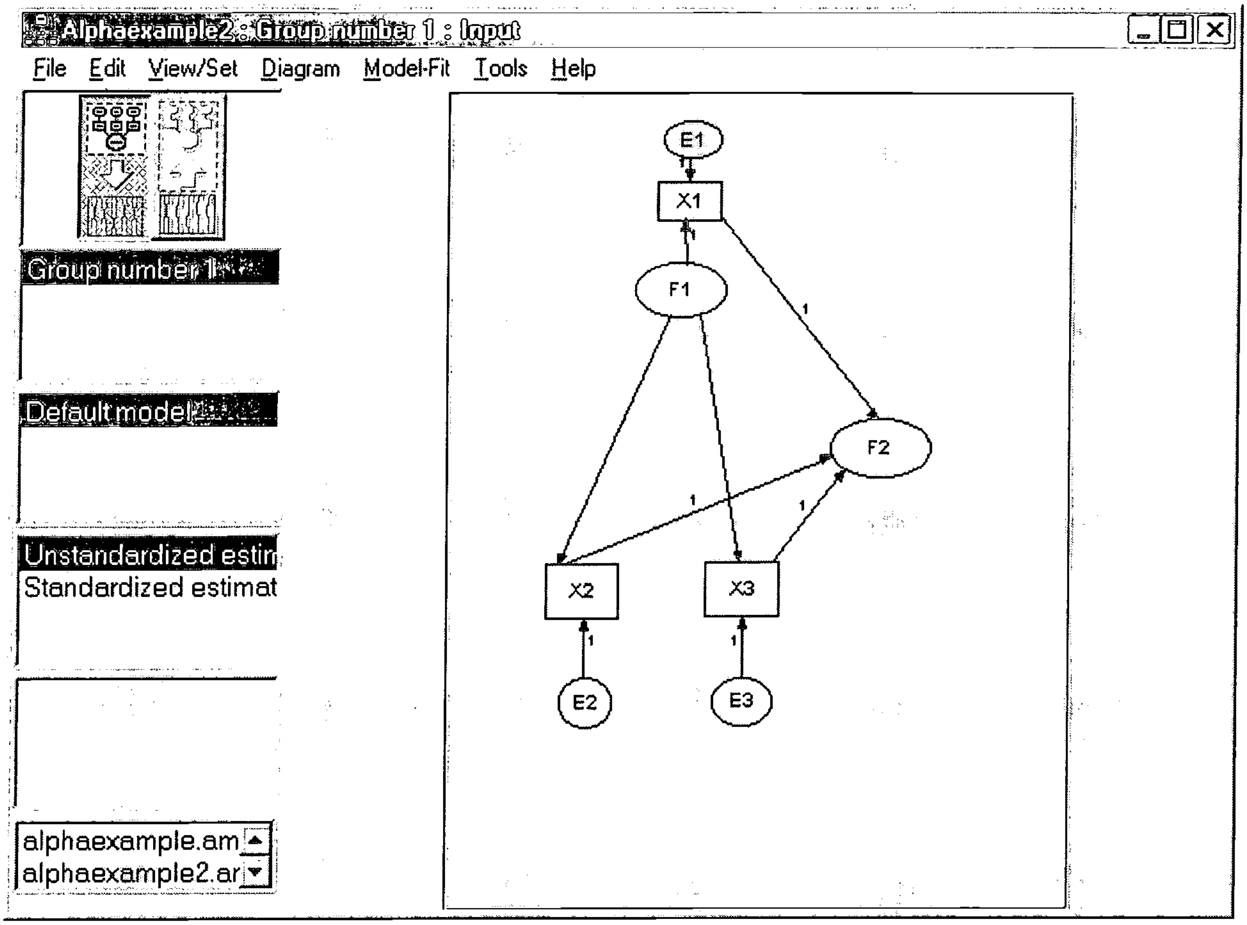

Figure 9 AMOS Model for Obtaining Measurement Reliability Estimate

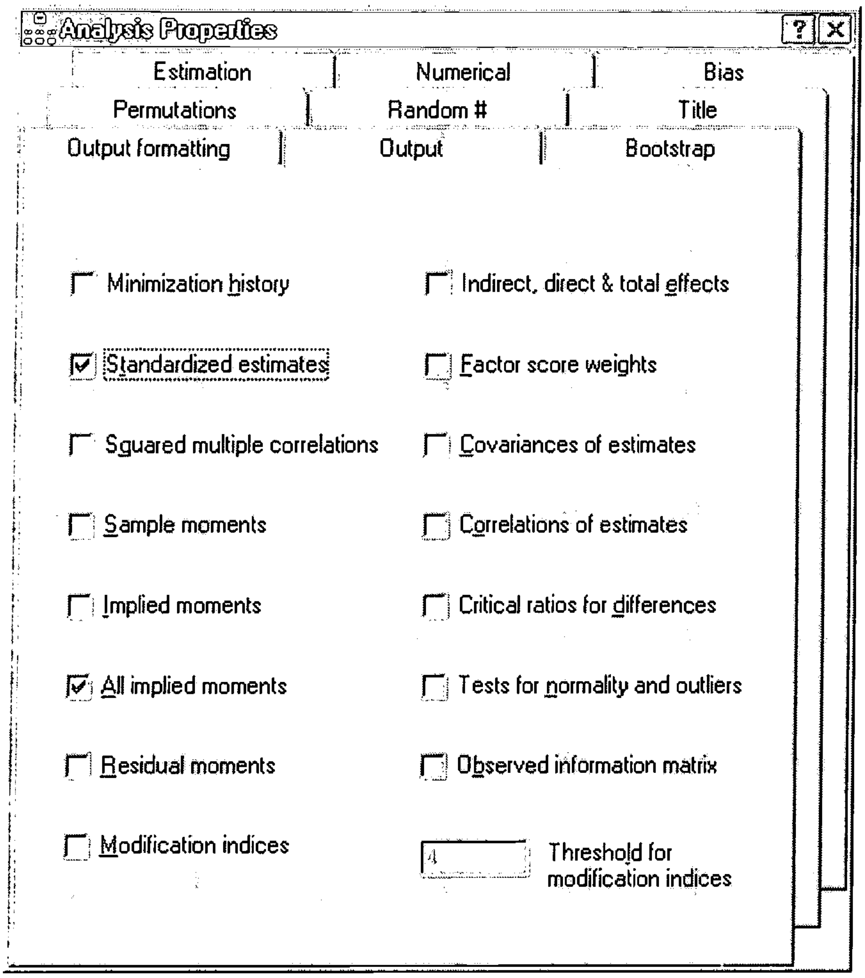

Figure 10 AMOS Output Option Selection for Obtaining Reliability Index

Signature:

Printed Name/Position/Title: Xitao Fan, Associate Professor

Organization/Address:
Curry School of Education
University of Virginia
PO Box 400277
Charlottesville, VA 22904-4277

Telephone: (804)243-8906

Fax: (804)924-1384

E-mail Address: xfan@virginia.edu

Date: May 29, 2001