Program Abstracts/Algorithms

A problem with the EQS V3.0 simulator

CONCEPCIÓN SAN LUIS, JUAN A. HERNÁNDEZ,
and ALFONSO SÁNCHEZ

University of La Laguna
La Laguna–Tenerife
Canary Islands, Spain

The structural equations program EQS V3.0 has proved to be an important tool for research in the field of Monte Carlo simulations. However, when used to replicate a simulation experiment, the simulator has been found to contain a serious flaw which makes it unusable for particularly complex models. This functional anomaly stems from the fact that, from a given replication onward, none of the remaining replications offer adequate and convergent solutions.

In recent years, structural equations models have proved to be an important tool for studying linear relations among variables. There has been a proliferation of articles on the robustness of the different solutions provided by structural equations systems, manipulating the influence exerted on the solutions by sample size, estimation method, normality of variables, and so forth, using as the basic tool the procedures known as Monte Carlo simulation (Bentler, 1983; Bentler & Bonnet, 1980; Browne, 1982; Huba & Harlow, 1986, 1987; La Du, 1986; MacCallum, 1986; Mulaik et al., 1989; Tanaka, 1987; Tanaka & Huba, 1985, 1987, 1989), among others.

Prior to the appearance of the EQS V3.0 program (Bentler, 1985), simulation work in the structural equations environment was a painstaking and repetitive task, given that it was necessary to generate the samples for each of the replications of each experimental condition one by one. With EQS V3.0, it became possible for the first time to generate and subsequently process all the experimental condition samples with one single program and one single application, thus allowing manipulation of all the factors that supposedly predicted the robustness and the correct estimation of the parameters of covariance structure models: estimation methods, distribution of observable variables as regards mean variance, skewness, and kurtosis. The author of the program recommended its use to facilitate implementation of all Monte Carlo simulation procedures (Chou, Bentler, & Satorra, 1991). How-

ever, when we used it to replicate a Monte Carlo simulation experiment, the simulator was found to contain a serious flaw which makes it unusable for particularly complex models. This functional anomaly stems from the fact that, from a given replication onward none of the remaining replications offer adequate and convergent solutions. Our hypothesis was that the EQS program used the last solution of parameters as starting values for the next replication and not the population starting values proposed.

In order to detect accurately the precise nature of the problem, a nonrecursive model was generated with eight variables and 14 *df*. Figure 1 gives the instructions for obtaining 100 replications of the matrix shown under the heading MATRIX, with the desired size and distribution. Population values were introduced as starting values, and these provided a perfect fit matrix ($\chi^2 = 0, p = 1$). Once the 100 replications were obtained, we looked at the seeds which provided the first inadequate and nonconvergent solution (Figure 2), and also the seed immediately before (adequate solution). From this point on, a new analysis with two replications was undertaken, taking as the initial seed the last one to generate an adequate solution (Replication 9) (Figure 3). Following this, a further analysis with one replication was undertaken, using the starting values obtained in the previous solution (Replication 9) and the seed which led to the first nonconvergent solution (Replication 10) (Figure 4).

Results

As shown in the first two columns of Figure 2, from Replication 10 on, all the solutions were nonconvergent or inadequate, or both. The solution to the replication of the processing of instructions in Figure 3, using the seed immediately prior to the nonconvergent matrix, was retained and subsequently used as the starting values for a further replication (Figure 4). In this case, the seed corresponded to Replication 10, which was that of the first erroneous replication (Figure 2). As shown in Figures 5A and 5B, the solution coincides with that obtained for Replication 10 in the first batch of processing of 100 replications (Figure 1). Last, we processed the same sequence of instructions with the random seed of Replication 10, although with population starting values. As Figure 5C shows, convergence and an adequate solution were achieved. Given that the cause of the problem was clear, we processed once again the program described in Figure 1, although this time requesting as resulting output the covariance matrix of each of the samples generated. This enabled us to process the 100 matrices independently, thus guaranteeing the use of population starting values rather than those obtained in the previous

Correspondence should be addressed to J. A. Hernández Cabrera, Facultad de Psicologia, Campus de Guajara, Universidad de La Laguna, La Laguna–Tenerife, Canary Islands, Spain (e-mail: jhdez@ull.es).

520

replication. In this way, 59% convergent and adequate solutions were obtained, and the adequate solutions were distributed randomly throughout the solutions (convergent and nonconvergent).

Discussion

Using the last solution of parameters as starting values for the next replication may be valid for simple normal distribution models. However, with more complex models such as nonrecursive models, featuring many variables and non-normal distribution, the processing of the generated sample in a given replication will lead to a solution far removed from the real one, and this will clearly diminish the likelihood of achieving convergence if the solution is used as the starting values for processing the next sample matrix.

Although the EQS has shown itself to be probably the best available application for the statistical treatment of covariance structure models, in view of its ductility and easy use, we feel that the problem we have detected renders use of the simulator impossible for the study of complex models. This is unfortunate, because its use with models of this type would prove to be one of the most interesting and promising aspects of the EQS, which in our opinion is highly efficient and of the highest quality. P. M. Bentler, one of the reviewers of this paper, pointed out to the present authors that this problem specifically occurred through the incorrect recovery of "beta" parameters from replication to replication. These are the parameters that represent the effects of dependent variables on other dependent variables. Consequently, if a simulation is done without a beta parameter, its results will be correct. Professor Bentler has notified us that the fault has now been corrected in the current production version of EQS/Windows 4.02.

```
/SPECIFICATION
  CASES=150; VARIA= 8; METHOD=ML;  ma=COR;
/simulation
  population=matrix;
  replication=100;
  tranformation=v1 to v8: s0.75, k1;
/EQUATIONS
  V1=    1F1 + E1;
  V2= 0.920*F1 + E2;
  V3=           1F2 + E3;
  V4=           0.961*F2 + E4;
  V5=                   0.512*F3 + E5;
  V6=                   0.485*F3 + E6;
  V7=                                 0.538*F4 + E7;
  V8=                                 0.604*F4 + E8;
  F1= 0.260*F2 + 0.346*F3 + 0.276*F4 + D1;
  F2= 0.438*F1 + 0.184*F3 + 0.346*F4 + D2;
/VARIAN
  F3= 1;
  F4= 1;
  E1 to E8= *;
  D1 TO D2= *;
/COVAR
  F3,F4= 0.254*;
/CONSTRAINT
  (F1,F3) = (F2,F4);
/output
 data='output1.rst';
/MATRIX
      1 0.716 0.672 0.646 0.281 0.266 0.282 0.317
  0.716     1 0.619 0.595 0.258 0.245  0.26 0.292
  0.672 0.619     1 0.762 0.262 0.248 0.335 0.376
  0.646 0.595 0.762     1 0.252 0.238 0.322 0.361
  0.281 0.258 0.262 0.252     1 0.248  0.07 0.078
  0.266 0.245 0.248 0.238 0.248     1 0.066 0.074
  0.282  0.26 0.335 0.322  0.07 0.066     1 0.325
  0.317 0.292 0.376 0.361 0.078 0.074 0.325     1
/end
```

Figure 1. Simulation with 100 replications of a nonrecursive model with 8 variables and 14 *df*.

0	0	484.58556	20.78318	14	.10732	.95711	.97029		
0	0	509.55593	8.93624	14	.83511	.98246	1.02103	SEED	449094354
0	0	341.96775	18.84463	14	.17097	.96914	.96914	SEED	2117302225
0	0	445.94682	14.94422	14	.38198	.96649	.99548	SEED	1836719449
0	0	549.08862	8.21605	14	.87777	.98504	1.02220	SEED	1353108009
0	0	428.95651	11.25618	14	.66580	.97376	1.01369	SEED	200382020
0	0	529.45031	7.93557	14	.89266	.98501	1.02419	SEED	559701348
0	0	501.86174	37.07672	14	.00072	.92612	.90260	SEED	88325120
0	0	355.27936	4.92480	14	.98683	.98614	1.05546	SEED	918272953
2	1	545.96142	27.75773	14	.01532	.94916	.94688	SEED	555010963
2	1	416.23231	16.54949	14	.28098	.96024	.98687	SEED	655816003
3	1	545.95726	26.48091	14	.02247	.95150	.95181	SEED	732932626
3	1	439.57768	22.22793	14	.07405	.94943	.96002	SEED	277273618
3	1	589.77856	40.75562	14	.00019	.93090	.90475	SEED	1090292732
2	1	489.68743	44.90251	14	.00004	.90830	.86613	SEED	1160547600
3	1	516.63158	13.77099	14	.46691	.97334	1.00094	SEED	1462772409
3	0	469.45617	16.90698	14	.26117	.96399	.98683	SEED	421340081
3	0	526.86469	20.90551	14	.10408	.96032	.97232	SEED	1293583912
9	1	436.95381	7.68019	14	.90531	.98242	1.03091	SEED	1480503517
2	1	542.31794	18.19613	14	.19799	.96645	.98368	SEED	1847813939
3	1	554.78112	32.68045	14	.00320	.94109	.92908	SEED	2013025619
3	1	519.29919	11.20913	14	.66953	.97841	1.01136	SEED	21962756
3	1	490.56476	21.74743	14	.08396	.95567	.96650	SEED	1759365059
3	1	388.25614	21.23910	14	.09566	.94530	.95981	SEED	941371112
4	1	447.20347	32.86002	14	.00302	.92652	.91002	SEED	339524930

Figure 2. Results of the first 25 replications with the random seeds corresponding to each replication. These results refer to 9 elements of model statistics: condition code (0 for normal condition), convergence (0 for model converged), null model chi-square, model chi-square, degrees of freedom, probability level, Bentler-Bonett normed fit index, Bentler-Bonett nonnormed fit index, and random seed.

```
/SPECIFICATION
  CASES=150; VARIA= 8; METHOD=ML;  ma=COR;
/simulation
  seed=918272953;
  population=matrix;
  replication=2;
  tranformation=v1 to v8: s0.75, k1;
/EQUATIONS
  V1=    1F1 + E1;
  V2= 0.920*F1 + E2;
  V3=             1F2 + E3;
  V4=          0.961*F2 + E4;
  V5=                   0.512*F3 + E5;
  V6=                   0.485*F3 + E6;
  V7=                            0.538*F4 + E7;
  V8=                            0.604*F4 + E8;
  F1= 0.260*F2 + 0.346*F3 + 0.276*F4 + D1;
  F2= 0.438*F1 + 0.184*F3 + 0.346*F4 + D2;
/VARIAN
  F3= 1;
  F4= 1;
  E1 to E8= *;
  D1 to D2= *;
/COVAR
  F3,F4= 0.254*;
/CONSTRAINT
  (F1,F3) = (F2,F4);
/output
  data='output2.rst' pa;
/MATRIX
      1 0.716 0.672 0.646 0.281 0.266 0.282 0.317
  0.716     1 0.619 0.595 0.258 0.245  0.26 0.292
  0.672 0.619     1 0.762 0.262 0.248 0.335 0.376
  0.646 0.595 0.762     1 0.252 0.238 0.322 0.361
  0.281 0.258 0.262 0.252     1 0.248  0.07 0.078
  0.266 0.245 0.248 0.238 0.248     1 0.066 0.074
  0.282  0.26 0.335 0.322  0.07 0.066     1 0.325
  0.317 0.292 0.376 0.361 0.078 0.074 0.325     1
/end
```

Figure 3. Program which requests two replications on the basis of the seed corresponding to Replication 9 (the last to converge). The aim is to determine the solution to this replication.

```
/SPECIFICATION
  CASES=150; VARIA= 8; METHOD=ML;  ma=COR;
/simulation
  SEED=555010963;
  population=matrix;
  replication=1;
  tranformation=v1 to v8: s0.75, k1;
/EQUATIONS
  V1=    1F1 + E1;
  V2= 0.923*F1 + E2;
  V3=            1F2 + E3;
  V4=            1.157*F2 + E4;
  V5=                      0.199*F3 + E5;
  V6=                      0.815*F3 + E6;
  V7=                                0.503*F4 + E7;
  V8=                                0.565*F4 + E8;
  F1=-1.455*F2 + 0.158*F3 + 1.250*F4 + D1;
  F2= 0.826*F1 + 0.030*F3 + 0.158*F4 + D2;
/VARIAN
  F3= 1;
  F4= 1;
  E1= 0.321*;
  E2= 0.345*;
  E3= 0.205*;
  E4= 0.164*;
  E5= 0.815*;
  E6= 0.280*;
  E7= 0.664*;
  E8= 0.667*;
  D1= 1.339*;
  D2= 0.218*;
/COVAR
  F3,F4= 0.283*;
/CONSTRAINT
  (F1,F3) = (F2,F4);
/output
 data='seed.rst' pa; co;
/MATRIX
     1 0.716 0.672 0.646 0.281 0.266 0.282 0.317
 0.716     1 0.619 0.595 0.258 0.245  0.26 0.292
 0.672 0.619     1 0.762 0.262 0.248 0.335 0.376
 0.646 0.595 0.762     1 0.252 0.238 0.322 0.361
 0.281 0.258 0.262 0.252     1 0.248  0.07 0.078
 0.266 0.245 0.248 0.238 0.248     1 0.066 0.074
 0.282  0.26 0.335 0.322  0.07 0.066     1 0.325
 0.317 0.292 0.376 0.361 0.078 0.074 0.325     1
/end
```

Figure 4. Program using seed corresponding to Replication 10, with the solution to the previous replication taken as the starting values.

A

```
2  1      545.96142        24.99535      14      .03461    .95422    .95754   1.00000
  .59159611D+00    .29352240D+00    .29522830D+00    .19463318D+00    .24027111D+00
  .40781696D+00    .41909370D+02    .72590331D+08    .48429077D+00    .20687019D+00
 -.54819782D+01    .40943693D+05   -.14500479D+02    .10527783D+01    .95528415D+00
  .10560786D+01    .94749576D+00    .63656767D+00
  .28638245D+00    .72131800D+00   -.14500479D+02
  .20670421D+02   -.44961394D+05
```

B

```
2  1      545.96142        24.99516      14      .03462    .95422    .95754   1.00000
  .59145179D+00    .29354406D+00    .29524341D+00    .19463169D+00    .24027248D+00
  .40780758D+00    .41037583D+02    .87023149D+08    .48427934D+00    .20686267D+00
 -.54247299D+01    .45045516D+05   -.14317837D+02    .10527867D+01    .95528265D+00
  .10560908D+01    .94749908D+00    .63656775D+00
  .28638238D+00    .72132466D+00   -.14317837D+02
  .20433613D+02   -.49457661D+05
```

C

```
0  0      545.96142        19.29198      14      .15409    .96466    .97957   1.00000
  .16783539D+00    .28627697D+00    .30284076D+00    .19263875D+00    .24208755D+00
  .40382304D+00    .18267505D+00    .11398937D+00    .66702236D+00    .32058040D+00
  .29309281D+00    .29156918D+00    .41913183D+00    .10442034D+01    .95330016D+00
  .84571381D+00    .88752225D+00    .63715009D+00
  .28536544D+00    .72408534D+00    .41913183D+00
  .27848078D+00    .42334432D+00
```

Figure 5. (A) Output resulting from processing of matrix generated from seed 555010963, with the results of the processing of the previous matrix from seed 918272953 used as the starting values. (B) Output of second replication obtained from first seed 918272953. Once this matrix is processed, EQS generates a new one from seed 555010963, giving the results that follow. (C) Output corresponding to seed 555010963, taking as starting values for processing of the matrix those which give rise to a matrix with perfect fit ($P = 1$), from which the random samples are taken.

REFERENCES

BENTLER, P. M. (1983). Some contributions to efficient statistics in structural models. *Psychometrika*, 48, 493-517.

BENTLER, P. M. (1985). *Theory and implementation of EQS, a structural equations program*. Los Angeles: BMDP Statistical Software.

BENTLER, P. M., & BONETT, A. G. (1980). Significance test and Goodness of fit in the analysis of variance structure. *Psychological Bulletin*, 88, 588-606.

BROWNE, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72-141). Cambridge: Cambridge University Press.

CHOU, C. P., BENTLER, P. M., & SATORRA, A. (1991). Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical & Statistical Psychology*, 44, 347-357.

HUBA, G. J., & HARLOW, L. (1986). Robust estimation for causal models: A comparison of methods in some developmental datasets. In P. B. Baltes, D. M. Featherman, & R. M. Lerner (Eds.), *Life-span developmental psychology* (Vol. 6, pp. 69-111). Hillsdale, NJ: Erlbaum.

HUBA, G. J., & HARLOW, L. (1987). Robust structural equation models: Implications for developmental psychology. *Child Development*, 58, 147-166.

LA DU, T. J. (1986). *Assessing the goodness of fit of linear structural equation models: The influence of sample size, estimation method and model specification*. Unpublished doctoral dissertation, New York University.

MACCALLUM, R. C. (1986). Specification searches in covariance structure models. *Psychological Bulletin*, 100, 107-120.

MULAIK, S. A., JAMES, L. R., VAN ALSTINE, J., BENNETT, N., LIND, S., & STILWELL, C. D. (1989). Evaluation of goodness of fit indexes for structural equation models. *Psychological Bulletin*, 105, 430-445.

TANAKA, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, 58, 134-146.

TANAKA, J. S., & HUBA, G. J. (1985). A fit index for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical & Statistical Psychology*, 38, 197-201.

TANAKA, J. S., & HUBA, G. J. (1987). Assessing the stability of depression in college students. *Multivariate Behavior Research*, 22, 5-19.

TANAKA, J. S., & HUBA, G. J. (1989). A general coefficient of determination for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical & Statistical Psychology*, 42, 233-239.