Inference in Latent Transition Models: Results from a Simulation Study

Stephanie L. Hyatt
and
Linda M. Collins

The Pennsylvania State University

The Methodology Center
Technical Report Series
#00-38

College of Health and Human Development

The Pennsylvania State University

Inference in Latent Transition Models:

Results from a Simulation Study

Stephanie L. Hyatt and Linda M. Collins

The Methodology Center

The Pennsylvania State University

Inference in Latent Transition Models:

Results from a Simulation Study

Latent transition analysis (LTA) is a method for estimating and testing stage-sequential models of development in longitudinal data. LTA uses an EM algorithm to obtain maximum likelihood estimates for all parameters. Because standard errors are not a byproduct of the EM algorithm, this method does not allow the user to conduct hypothesis tests. Recently, a data augmentation procedure has been added to LTA; this procedure uses the data, the LTA model, and the EM estimates to multiply impute the latent variables. The within-imputation and between-imputation variability can then be combined to yield an overall estimate of the standard error for each parameter. This procedure provides information about the variability of point estimates and enables the user to conduct hypothesis tests.

The purpose of this technical report is to record how parameter estimates and confidence intervals based on data augmentation compare to known parameter values for a particular latent transition model. A simulation was conducted in order to assess the bias of parameter estimates, the bias of their associated standard errors, and the coverage of the associated 95% confidence intervals. We investigated conditions under which the data augmentation procedure can successfully recover parameters by varying sample size, the strength of the measurement parameters, the presence of missing data, and the number of imputations.

Latent Transition Analysis

Latent class analysis (Goodman, 1974) uses the EM algorithm to fit latent class models to data, where a categorical latent variable is indicated by multiple categorical manifest items. This

method yields estimates of the probability of class membership, as well as the amount of measurement error associated with each indicator.

Latent transition analysis (LTA; Collins & Wugalter, 1992) extends latent class theory to encompass multiple times of measurement, allowing the estimation of stage-sequential models of development in longitudinal data. The prevalence of stages and incidence of stage transitions over time are estimated using multiple categorical indicators of the latent dynamic variable measured at each time. Consider for the purposes of this study a model for the onset of adolescent substance use for subjects measured in 9th and 10th grade. We could model the onset process by following adolescents' experimentation with alcohol, cigarettes, drunkenness, and marijuana across these two years. A model of onset does not require all subjects to advance to the most advanced stage; it is merely assumed that if they do, they follow a specified path. The following presentation of the LTA mathematical model and the model used in the simulation involve only a dynamic latent variable. LTA can, however, include both a dynamic and a static latent variable (see Collins & Wugalter, 1992 for a presentation of this more general model).

The LTA Mathematical Model

To aid explanation, the LTA mathematical model will be presented in terms of two times of measurement and five manifest indicators of the latent status at each time. The model can be directly extended to other latent transition problems. Let the first time of measurement be Time t and the second be Time t+1. Suppose also that the five manifest indicators of latent status are Item 1 with i, $i' = 1, ..., I$ response categories, Item 2 with j, $j' = 1, ..., J$ response categories, and so on, where i, j, k, l, and m refer to item responses at Time t and i', j', k', l', and m' refer to item responses at Time t+1. Let $p = 1, ..., S$ denote the latent status at Time t and $q = 1, ..., S$ the latent

status at Time t+1. A response can then be represented by $Y = \{i,j,k,l,m,i',j',k',l',m'\}$, the vector of responses to the five latent status indicators at Time t and the five latent status indicators at Time t+1.

There are three sets of parameters estimated in this model: δ parameters, τ parameters, and ρ parameters. The estimated probability of having a particular response pattern is:

$$P(Y) = \sum_{p=1}^{S} \sum_{q=1}^{S} \delta_p \rho_{i|p} \rho_{j|p} \rho_{k|p} \rho_{l|p} \rho_{m|p} \tau_{q|p} \rho_{i'|q} \rho_{j'|q} \rho_{k'|q} \rho_{l'|q} \rho_{m'|q}$$

where

δ_p is the probability of being in latent status p;

$\rho_{i|p}$ is the probability of response i to Item 1 at Time t, given membership in latent status p at Time t (similarly, $\rho_{i'|q}$ represents the probability of response i' to Item 1 at Time t+1, given membership in latent status q at Time t+1); and

$\tau_{q|p}$ is the probability of membership in latent status q at Time t+1 given membership in latent status p at Time t.

Data Augmentation and Multiple Imputation

The analyses presented here were performed using WinLTA 3.0, a new version of the LTA program which includes the capability of estimating standard errors for each parameter. The EM algorithm in LTA has been tested and applied extensively, but this study is the first to investigate the performance of the new data augmentation feature of LTA.

Multiple imputation is a general approach to missing data problems that has been shown to produce high quality estimates and reliable standard errors (Schafer, 1997). This approach to missing data has been incorporated in the WinLTA program, where the latent variables are

treated as missing data and multiply imputed. Multiple imputation employs data augmentation, an iterative simulation procedure similar to the EM algorithm. It is important to note that the current version of WinLTA uses a Jeffreys prior; the choice of the prior used in data augmentation may have a strong impact on the results when the data supply little information. This iterative procedure produces a sequence of plausible sets of LTA parameter estimates. Sets spaced sufficiently apart, rather than consecutive ones, tend to represent independent draws of parameters. Typically draws are spaced far enough apart if this distance equals the number of iterations within which EM converged to the maximum likelihood estimates. Multiple draws of LTA parameters, or imputations, are retained (five or ten draws, in this study) and analyzed as if the latent variables are known. Parameter estimates and their standard errors are retained from each imputation's analysis, and are combined according to rules defined by Rubin (1987). The resultant parameter estimates and their standard errors can then be used to describe the variability about the parameter estimates and to conduct hypothesis tests. (See Schafer, 1997 and Brunner & Schafer, 1997 for more details on multiple imputation and its application in LTA.)

Methods

Overview

The purpose of this simulation study is to investigate the bias of parameter estimates and standard errors, and the coverage of confidence intervals based on LTA's data augmentation algorithm. Coverage refers to the frequency that a parameter's confidence interval contains the true parameter. Performance related to selected parameters will be examined to see if LTA's data augmentation algorithm successfully recovers parameters in the long run. Datasets were

produced based on LTA models and known parameters, thus giving us the ability to compare parameter estimates from generated random data against the true parameter values.

The Model

The model in this study was based on several recent empirical studies using LTA to examine factors relating to adolescent substance use onset (see Hyatt & Collins, 1999). Results from these empirical studies were modified slightly (an older cohort was used and the grouping variable was eliminated) and used to define the model for this simulation study. This model involves stage-sequential development over two times of measurement. There are five dichotomous indicators of the latent status, and the model includes eight latent statuses. A substantive example of this model might be substance use measured in 9th grade and again in 10th grade, with the following five manifest items: lifetime alcohol use, lifetime cigarette use, two measures of drunkenness in the past year, and lifetime marijuana use. Figure 1 depicts the model used in this study, which includes the following eight latent statuses defined from the above five items:

1) No use

2) Alcohol use

3) Cigarette use

4) Alcohol use, cigarette use

5) Cigarette use, marijuana use

6) Alcohol use, cigarette use, drunkenness

7) Alcohol use, cigarette use, marijuana use

8) Alcohol use, cigarette use, drunkenness, marijuana use

Table 1 displays the parameters used to generate the data. The δ parameters represent the probability of membership in each latent status at Time 1. The τ parameters represent the probability of transitioning to each latent status at Time 2 given latent status membership at Time 1. The ρ parameters represent the probability of answering 'No' to an item given latent status membership. There are two sets of ρ parameters shown in Table 1. The "strong" condition involved ρ parameters equal to .1 and .9, and the "weak" condition involved values of .3 and .7.

Data generation

The procedure for generating random data in this study is identical to the method used in Hyatt and Collins (1998). See Flaherty and Collins (1998) for details on the program used to generate response-format data. For each combination of the factors listed below, 1000 random multivariate datasets were generated from a population with known parameters. For the conditions involving missing data, the datasets from the corresponding cell with no missing data were degraded. For the conditions involving 5 imputed datasets, the first 5 imputations from the corresponding cell with 10 imputations were used. Each dataset was used in the initial EM analysis, and retained for use with data augmentation if it converged to maximum likelihood estimates within 5000 imputations, was identified, and did not have the "naming problem"; these three criteria are explained in the analysis section below. Additional datasets were then generated in each condition so that the total number of usable datasets per cell totaled 1000. The average bias, average standard error, and coverage were calculated for each parameter estimate by averaging across these final 1000 datasets.

Four factors were varied in this study. First, the *sample size* was varied such that N=300 and N=1000; we expected estimates based on larger samples to have less bias. Second, for each

sample size, two sets of *rho parameters* were used to generate data. The first set employed strong measurement parameters of .1 and .9, and the second set employed weak measurement parameters of .3 and .7. Weak ρ parameters are an indication of higher levels of measurement error, which is known to make the recovery of true parameter values more difficult. The third factor varied in this study is the *presence of missing data.* The two conditions were no missing data or missing data, where approximately 12% of the data were missing completely at random due to a combination of spot missing and attrition. Attrition occurs when a subject has data at Time 1 but not Time 2, and spot missing refers to data not being available for a particular item. It was hypothesized that missing data will introduce some level of bias in parameter estimates. The final factor varied in this study is the *number of imputed datasets*; the two conditions were 5 and 10 imputations. In theory, a larger number of imputed datasets should increase the accuracy of the standard errors, however 5 imputations is expected to be quite sufficient; therefore the gains from doubling the number of imputations (and therefore the time of computation) to 10 are expected to be small. Table 2 displays the experimental design used in this simulation, showing all combinations of the factors which were examined.

Analysis

Overview of the Analysis. In order to assess the performance of data augmentation for estimating standard errors of LTA parameter estimates, 1000 datasets were randomly generated from a known model in each combination of factors. Each dataset was then estimated via the EM algorithm using two unique sets of starting values. Identification of each dataset was assessed, and additional datasets were generated in order to obtain 1000 identified datasets for each cell of the simulation. This allows for comparisons to be made across cells with equivalent

statistical power. These datasets were then analyzed using data augmentation, and ten imputations for each dataset were produced and analyzed. Figure 2 presents a flowchart of this process.

Details about the EM Estimation. When estimating LTA models, various types of parameter restrictions can be imposed on parameters in order to ensure identifiability of the model and aid in interpretation. For the purposes of this study, the δ parameters and all logical τ parameters were estimated freely. Table 3 shows which τ parameters were estimated freely and which were fixed to zero. To aid in model identification, a pattern of constraints was imposed on the ρ parameters. First, we constrained the ρ parameters to be equal across time, which is necessary for making the interpretation of the transition probabilities meaningful. Second, additional constraints were added so that only two ρ parameters were estimated for each of the five items. Table 4 displays the constraint pattern used for the estimation of the ρ parameters in this study. This pattern is substantially more parsimonious than estimating all ρ parameters freely. Each dataset was analyzed using two unique sets of starting values. Convergence of the EM algorithm was reached when a mean absolute deviation between successive sets of parameter estimates smaller than .00001 was reached within 5000 iterations. For each dataset, the G^2 and the final parameter estimates based on the two sets of starting values were compared to determine if the solution was identified. Datasets were considered identified when the G^2 values were equal out to at least one decimal place and no parameter estimate differed by more than .03.

Tables 5 displays information about the original 1000 datasets generated under each condition. For each combination of factors 1000 datasets were produced. The first column gives unique designations for each cell in the design (this is primarily for our records), with extensions

'a' and 'b' referring to the two different sets of starting values. The columns labeled 'N' and 'Rho Strength' refer to the sample size and strength of the rho parameters used to generate each group of 1000 datasets (strong rho parameters were .9 and .1, and weak rho parameters were .7 and .3). The 'Missing Data' column refers to whether or not missing data was imposed on the datasets. The 'Start Values' column displays which set of starting values was used, with close starting values being nearer the true parameters.

The '# Converged' column lists the number of datasets out of 1000 that converged within 5000 iterations. In this study all datasets in each cell converged within 5000 iterations. The number of identified datasets appears in the '# Identified' column. The '# With Naming Problem' column shows the number of identified datasets which had two or more latent statuses that were not distinguishable due to similar patterns of the ρ parameters. In this study datasets were discarded if the eight latent statuses were not distinguishable or if the order of the latent statuses did not correspond to the order in the true model. The naming problem did not occur under any conditions when ρ parameters were strong (values of .9 and .1). When the ρ parameters were weak, however, as many as 9% of identified datasets had to be discarded. The column labeled '# Good Sets' shows the total number of identified datasets without the "naming problem" out of 1000. Additional datasets were then generated in order to have a total of 1000 that will be usable for DA. The final column in Table 5, 'Total Sets Generated,' shows the number of datasets that were required in order to produce 1000 usable datasets.

The number of the original 1000 datasets that were usable varied substantially across cells. When strong ρ parameters were used to generate data, at least 98% of the datasets were identified and did not have the naming problem. With weak ρ parameters, however, many more

datasets were not identified or had naming problems. Less than 70% of the datasets involving a small sample size, weak ρ parameters, and missing data were usable.

The parameter estimates based on the EM algorithm were used as starting values for data augmentation. DA also requires that the user specify *k*, the number of iterations between each imputation. A good rule of thumb is to set this equal to the number of iterations the EM algorithm required to converge for a particular dataset. This rule was followed in this simulation, so that the value of *k* for each dataset was determined by how quickly the EM converged for that dataset. Table 6 shows the mean number of iterations, i.e. the mean value of *k,* for each cell, along with the standard deviation, minimum, and maximum value of *k* within each cell. Clearly datasets with weak ρ parameters require data augmentation to run much longer between imputations on average. Sample size appears to have little effect on the average number of iterations to convergence. The presence of missing data results in a small increase in the number of iterations to convergence on average. Parameter estimates and their variance estimates, along with 95% confidence intervals, were retained from the data augmentation runs.

Results

Table 7 contains information on five selected parameter estimates under all conditions. The bolded parameters in Table 1 correspond to the parameters that are reported in Table 7. These parameters include two δ parameters: Delta1, the probability of membership in the 'No use' latent status (the true value is 0.4), and Delta8, the probability of membership in the 'Alcohol use, cigarette use, drunkenness, and marijuana use' latent status (the true value is 0.1). Two τ parameters are examined: Tau11, the probability of 'No use' at Time 2 given 'No use' at

Time 1 (the true value is 0.7), and Tau46, the probability of 'Alcohol use, cigarette use, and drunkenness' at Time 2 given 'Alcohol use and cigarette use' at Time 1 (the true value is 0.1). Table 7 also reports the results for Rho15, the ρ parameter representing the probability of responding 'No' to the marijuana use item, given membership in the 'No use' latent status (the true value is 0.9 for the strong measurement condition and 0.7 for the weak measurement condition). Note that this parameter was constrained to be equal across time.

There is an obvious omission in Table 7. The cells that represent weak rhos, large sample size, and missing data are not reported because the analysis was deemed impractical. Each dataset runs for so many iterations, and each iteration requires so much processor time, that computational aspects lead us to recommend that data augmentation not be used under these conditions.

Table 7 allows us to examine which factors affect the unbiased recovery of true parameter values. There are six entries in each cell of the table. The first entry, 'N', represents the number of the 1000 usable datasets which produced rho parameters in the expected range for all ten imputations. This criterion will be discussed in more detail below. All other entries in the cells of Table 7 are based on their corresponding value of N reported here. The entry called 'Est' refers to the average parameter estimate based on data augmentation, or the average of the $\overline{Q}$'s. 'Bias' is simply the difference between the average parameter estimate and the known true value of the parameter. It is desirable to have a bias of 0, although when there is not much information we expect that the prior information may introduce bias towards the values of the Jeffreys prior. Hypothesis tests were conducted in each cell to test the null hypothesis that the bias equals 0. Results were bolded to indicate one of two things: either the bias was less than .02, which was

deemed to be insignificant by the investigators, or the confidence interval around the bias contains 0, and thus bias was not statistically significant (with 95% certainty). Therefore, we will consider bolded results to represent conditions under which parameters were recovered well, and the bias was small enough to be considered acceptable. The entry called 'Var(Est)' refers to the sample variance of the N average parameter estimates. 'Avg. T' corresponds to the average total variance estimate across the N datasets. In the long run we expect the average total variance to be approximately equal to the sample variance of the parameter estimate. Finally, 'Coverage' refers to the percentage of times the 95% confidence interval based on data augmentation contains the parameter's true value. As a point of comparison, Table 7 also reports the average maximum-likelihood parameter estimate and the sample variance for all 1000 datasets.

The simulation was designed so that each set of conditions would have a sample of 1000 datasets so that comparisons across cells would be based on the same amount of information. The multiple imputation approach to obtaining standard error estimates for each parameter relies on taking multiple draws from the predictive posterior distribution of each parameter. An interesting dilemma developed when we found that the distribution of individual rho parameters covered 0.5. Table 1 shows how the eight latent statuses are defined. For example, Status 1 (representing the 'No use' latent status) is defined by a low probability of responding 'Yes' to each substance use item. Because all delta and tau parameters in the LTA model are defined by the rho parameters, it is unclear how to interpret these parameters when the rho parameters drawn for each item have an unexpected pattern. Further complicating matters, because ten multiple imputations were drawn in the analysis of each dataset, patterns of rho parameters could vary across the imputations for a single dataset. In order to have interpretable parameters which

correspond to this latent transition model, datasets were included in the analysis for the results in Table 7 only if all 10 imputations produced estimates for the rho parameters which were consistent with our model. This selection seriously restricts the data informing results under the weak rho parameters condition; with a sample size of 300, only 76 out of 1000 datasets were analyzed in the no missing data condition, and 38 out of 1000 datasets were used in the missing data condition. The estimates of the parameters and their standard errors may be biased by this selection effect, and therefore results based on conditions with weak rhos should be interpreted with great caution. Conditions with strong rhos, however, can be examined thoroughly. Because of this, the discussion of the results will focus on conditions containing strong rho parameters.

δ Parameters

For this model δ parameters are the unconditional probability of being in a certain latent status at Time 1. In this study, Delta1 has a value of .4 (see Table 1), indicating that subjects have a 40% chance of being in the 'No Use' status at Time 1. Under the strong ρ parameter condition, this parameter was estimated without bias regardless of the sample size or the presence of missing data. The coverage of the confidence intervals is excellent under all strong ρ conditions, ranging from .95 to .96 in every case.

Delta8 has a known value of .1, representing a 10% chance of subjects being in the most advanced latent status at Time 1. With strong ρ parameters, this parameter was estimated without bias regardless of the sample size or the presence of missing data. Coverage of the confidence intervals ranges from .91 to .95 for all strong measurement conditions, with the best coverage in the cells relating to strong ρ parameters, large sample size, and no missing data.

τ Parameters

In this study, τ parameters represent the probability of being in a certain latent status at Time 2 given membership in a certain latent status at Time 1. Tau11, the probability of being in the 'No Use' status at both Time 1 and Time 2, has a known value of .7 (see Table 1). This parameter was estimated without significant bias under all conditions involving strong ρ parameters. Interestingly, the bias was smallest in the cells corresponding to strong measurement, small sample size, and missing data present. Coverage ranged from .94 to .98, ensuring reasonable coverage under any set of conditions with strong measurement.

Tau46 represents the probability of being in the 'Alcohol and Cigarettes' status at Time 1 and the 'Alcohol, Cigarettes, and Drunk' status at Time 2. The true value of this parameter is .1, and it was successfully recovered with strong ρ parameters and a large sample size. The combination of strong measurement and a large sample size ensured a bias of less than .005. In all strong measurement conditions coverage was .97 or higher. Under conditions involving strong measurement and a small sample size, bias ranges from .02 to .04.

ρ Parameter

In this study, ρ parameters represent the probability of responding correctly or incorrectly to a dichotomous item given membership in a latent status. Rho15 corresponds to the probability of answering 'No' to the Marijuana item given membership in the 'No Use' latent status. This parameter was estimated without bias under all conditions characterized by strong measurement. Coverage of the confidence intervals was .95 or higher with the large sample size, and .94 with the small sample size.

Effect of the Number of Imputed Datasets

The cost of imputing ten datasets rather than five lies mainly in processor time. This study clearly shows that there is no substantial gain in doubling the number of imputed datasets. Within any combination of ρ strength, sample size and missing data, the results are very similar for 5 versus 10 imputations. In every case when comparing the two choices for the number of imputations, when a parameter was estimated without bias for 5 imputations, the bias was also nonsignificant for 10 imputations. There is no evidence that increasing the number of imputations yields less bias or better coverage, and thus we recommend that five imputations is sufficient for similar LTA models.

Discussion and Conclusions

The results of this study suggest that the use of data augmentation to estimate standard errors of parameters for LTA models is appropriate and useful when the ρ parameters are strong. Small sample sizes, missing data, and weak measurement parameters are conditions which reflect a lack of information, whereas large sample sizes, no missing data, and strong measurement parameters are conditions which indicate that there is more available information. Under poor conditions there many not be enough information to apply data augmentation with confidence.

In general, parameter estimates based on data augmentation are more biased than the maximum-likelihood estimates. However, conditions involving strong ρ parameters usually yield results that are not substantially biased and that have reasonable coverage. A larger sample size also appears to produce slightly less biased parameter estimates and better coverage. The presence of missing data did not appear to reduce the performance of data augmentation

substantially. Increasing the number of imputations from 5 to 10 did not substantially improve the performance of data augmentation in LTA.

We know from past research (eg. Hyatt & Collins, 1998) that the strength of the ρ parameters tends to be the most important factor in the estimation of LTA models. We believe that this is also the most important factor in the performance of data augmentation in LTA. However, in this study it is difficult to assess the effect of the ρ parameters because we eliminated datasets when one or more of the ten imputed datasets produced a pattern of rho's inconsistent with the true model.

At this time we can confidently recommend the use of data augmentation for obtaining estimates of standard errors in latent transition models involving strong measurement parameters. In order to have the highest level of confidence when estimating standard errors in latent transition models, one should have strong ρ parameters and a large sample size. Even with the smaller sample size, bias was small and coverage was adequate in most cases when the ρ parameters were strong. Until more research is conducted on the use of data augmentation with weak ρ parameters, however, we recommend that the procedure not be used when the measurement is not strong. Users can base this decision on the results provided by the EM algorithm. One check that users can perform to ensure that results based on data augmentation are clearly interpretable is to examine the pattern of the ρ parameters from each imputation and verify that they are consistent.

Future research on the use of data augmentation to obtain multiple imputations in LTA models may focus on model complexity, model correctness (specification of model), and sparseness of cells in the LTA model. More importantly, however, research is needed in order to

better understand how to interpret results when rho parameters do not hold consistent patterns across the multiple imputations. One promising approach will be to apply an informative prior other than the Jeffreys prior in the LTA model.

References

Brunner, R. M., & Schafer, J. L. (1997). Inference for latent transition models (Technical Report Number 97-11). University Park, PA: The Pennsylvania State University, The Methodology Center.

Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. Multivariate Behavioral Research, 27, 131-157.

Flaherty, B. P., & Collins, L. M. (1998). The LTA data generator: Its care and use (Technical Report Number 98-19). University Park, PA: The Pennsylvania State University, The Methodology Center.

Hyatt, S. L., & Collins, L. M. (1999). Pubertal timing and the stages of drug use. (Technical Report Number 99-35). University Park , PA: The Pennsylvania State University.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika, 61, 215-231.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. New York, NY: J. Wiley & Sons.

Schafer, J. L. (1997). Analysis of incomplete multivariate data. London: Chapman & Hall.

Table 1

Parameter Values for Population

δ Parameters at Time 1

Status 1 No Use	Status 2 A	Status 3 C	Status 4 AC	Status 5 CM	Status 6 ACD	Status 7 ACM	Status 8 ACDM
.40	.10	.15	.10	.05	.05	.05	**.10**

τ Parameters (Transition Probabilities)

Time 2: Time 1:	Status 1	Status 2	Status 3	Status 4	Status 5	Status 6	Status 7	Status 8
Status 1	**.7**	.1	.1	.02	.02	.02	.02	.02
Status 2	.0	.7	.0	.15	.0	.05	.05	.05
Status 3	.0	.0	.5	.3	.05	.05	.05	.05
Status 4	.0	.0	.0	.6	.0	**.1**	.1	.2
Status 5	.0	.0	.0	.0	.5	.0	.3	.2
Status 6	.0	.0	.0	.5	.0	.3	.0	.2
Status 7	.0	.0	.0	.0	.0	.0	.7	.3
Status 8	.0	.0	.0	.0	.0	.0	.3	.7

ρ Parameters (Measurement Parameters) for 'No' Response

	Item 1 Alcohol	Item 2 Cigarettes	Item 3 Drinks	Item 4 Drunk	Item 5 Marijuana
Status 1	.9 / .7[1]	.9 / .7	.9 / .7	.9 / .7	**.9 / .7**
Status 2	.1 / .3	.9 / .7	.9 / .7	.9 / .7	.9 / .7
Status 3	.9 / .7	.1 / .3	.9 / .7	.9 / .7	.9 / .7
Status 4	.1 / .3	.1 / .3	.9 / .7	.9 / .7	.9 / .7
Status 5	.9 / .7	.1 / .3	.9 / .7	.9 / .7	.1 / .3
Status 6	.1 / .3	.1 / .3	.1 / .3	.1 / .3	.9 / .7
Status 7	.1 / .3	.1 / .3	.9 / .7	.9 / .7	.1 / .3
Status 8	.1 / .3	.1 / .3	.1 / .3	.1 / .3	.1 / .3

[1] Values of .9 and .1 correspond to strong measurement condition.
Values of .7 and .3 correspond to weak measurement condition.

Table 2

Experimental Design: Number of Datasets Generated per Condition

	Weak ρ parameters (.3, .7)				Strong ρ parameters (.1, .9)			
	300 Subjects		1000 Subjects		300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data	No Missing Data	Missing Data	No Missing Data	Missing Data
5 Imputed Datasets	1000	1000	1000	1000	1000	1000	1000	1000
10 Imputed Datasets	1000	1000	1000	1000	1000	1000	1000	1000

Table 3

τ Parameter Restrictions

	Status 1	Status 2	Status 3	Status 4	Status 5	Status 6	Status 7	Status 8
Status 1	FR[1]	FR	FR	FR	FR	FR	FR	FR
Status 2	0	FR	0	FR	0	FR	FR	FR
Status 3	0	0	FR	FR	FR	FR	FR	FR
Status 4	0	0	0	FR	0	FR	FR	FR
Status 5	0	0	0	0	FR	0	FR	FR
Status 6	0	0	0	FR	0	FR	0	FR
Status 7	0	0	0	0	0	0	FR	FR
Status 8	0	0	0	0	0	0	FR	FR

[1] FR represents parameters freely estimated.
0 represents parameters fixed to zero.

Table 4

ρ Parameter Restrictions

	Alcohol[1]	Cigarettes	5+ Drinks	Drunk	Marijuana
Latent Status 1	a	c	e	g	i
Latent Status 2	b	c	e	g	i
Latent Status 3	a	d	e	g	i
Latent Status 4	b	d	e	g	i
Latent Status 5	a	d	e	g	j
Latent Status 6	b	d	f	h	i
Latent Status 7	b	d	e	g	j
Latent Status 8	b	d	f	h	j

[1] Parameters denoted by the same letter are constrained equal.

Table 5

Simulation Results for EM Estimation

Dataset Desig-nation	N	Rho Strength	Missing Data	Start Values	# Con-verged	# Iden-tified	# With Naming Problem	# Good Sets	Total Sets Generated
dasim1a dasim1b	300	weak	No	close far	1000 1000	823	54	769	1283
dasim2a dasim2b	300	weak	Yes	close far	1000 1000	764	72	692	1449
dasim3a dasim3b	300	strong	No	close far	1000 1000	993	0	993	1007
dasim4a dasim4b	300	strong	Yes	close far	1000 1000	993	0	993	1007
dasim5a dasim5b	1000	weak	No	close far	1000 1000	885	1	884	1132
dasim6a dasim6b	1000	weak	Yes	close far	1000 1000	832	1	831	N/A
dasim7a dasim7b	1000	strong	No	close far	1000 1000	980	0	980	1020
dasim8a dasim8b	1000	strong	Yes	close far	1000 1000	982	0	982	1018

Identification Criteria: G^2 equal to one decimal place and no estimate difference greater than .03.

Table 6

Summary of Iterations used for DA Estimation

	N	Rho Strength	Missing Data	Mean Iterations	SD Iterations	Min Iterations	Max Iterations
dasim1	300	weak	0%	623	356	138	4051
dasim2	300	weak	12%	670	382	208	3620
dasim3	300	strong	0%	67	24	29	476
dasim4	300	strong	12%	84	27	38	273
dasim5	1000	weak	0%	647	217	286	2407
dasim6	1000	weak	12%	740	263	260	2856
dasim7	1000	strong	0%	54	11	30	117
dasim8	1000	strong	12%	70	14	41	136

Table 7

Simulation Results for Selected Parameters

DELTA1: No Use (δ = .4)

Weak rho parameters (.3 and .7):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	N = 76 Est = .4311 Bias = .0311 Var(Est) = .0095 Avg. T = .0206 Coverage = .961	**N = 38** **Est = .4145** **Bias = .0145** **Var(Est) = .0111** **Avg. T = .0218** **Coverage = .974**	N = 676 Est = .4576 Bias = .0576 Var(Est) = .0076 Avg. T = .0115 Coverage = .942	
DA estimate: 10 Imputed Datasets	N = 76 Est = .4373 Bias = .0373 Var(Est) = .0061 Avg. T = .0191 Coverage = .961	**N = 38** **Est = .4236** **Bias = .0236** **Var(Est) = .0084** **Avg. T = .0214** **Coverage = 1.00**	N = 676 Est = .4548 Bias = .0548 Var(Est) = .0067 Avg. T = .0101 Coverage = .942	
ML Estimate	Estimate = .3669 Var = .0283	Estimate = .3526 Var = .0316	Estimate = .3859 Var = .0094	

Strong rho parameters (.1 and .9):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	**N = 1000** **Est = .4029** **Bias = .0029** **Var(Est) = .0018** **Avg. T = .0020** **Coverage = .959**	**N = 990** **Est = .4112** **Bias = .0112** **Var(Est) = .0022** **Avg. T = .0023** **Coverage = .954**	**N = 1000** **Est = .4000** **Bias = .0000** **Var(Est) = .0005** **Avg. T = .0005** **Coverage = .952**	**N = 1000** **Est = .3998** **Bias = -.0002** **Var(Est) = .0059** **Avg. T = .0006** **Coverage = .952**
DA estimate: 10 Imputed Datasets	**N = 1000** **Est = .4041** **Bias = .0041** **Var(Est) = .0017** **Avg. T = .0019** **Coverage = .953**	**N = 990** **Est = .4103** **Bias = .0103** **Var(Est) = .0021** **Avg. T = .0022** **Coverage = .956**	**N = 1000** **Est = .3984** **Bias = -.0016** **Var(Est) = .0004** **Avg. T = .0005** **Coverage = .958**	**N = 1000** **Est = .3999** **Bias = -.0001** **Var(Est) = .0005** **Avg. T = .0006** **Coverage = .958**
ML Estimate	Estimate = .3998 Var = .0017	Estimate = .3993 Var = .0020	Estimate = .4001 Var = .0004	Estimate = .3998 .0005

Note: Bolded results indicate that bias is not significantly different from 0 with 95% confidence or bias is ≤ .02.

Table 7 (continued)

Simulation Results for Selected Parameters

DELTA8: Alcohol+Cigarettes+Drunk+Marijuana (δ = .1)

Weak rho parameters (.3 and .7):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	N = 76 Est = .1411 Bias = .0411 Var(Est) = .0025 Avg. T = .0054 Coverage = .934	N = 38 Est = .1407 Bias = .0407 Var(Est) = .0025 Avg. T = .0051 Coverage = .895	N = 676 Est = .1331 Bias = .0331 Var(Est) = .0017 Avg. T = .0024 Coverage = .922	
DA estimate: 10 Imputed Datasets	N = 76 Est = .1370 Bias = .0370 Var(Est) = .0021 Avg. T = .0046 Coverage = .974	N = 38 Est = .1420 Bias = .0420 Var(Est) = .0022 Avg. T = .0047 Coverage = .947	N = 676 Est = .1344 Bias = .0344 Var(Est) = .0016 Avg. T = .0023 Coverage = .920	
ML Estimate	Estimate = .0705 Var = .0038	Estimate = .0726 Var = .0042	Estimate = .0864 Var = .0019	

Strong rho parameters (.1 and .9):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	**N = 1000** **Est = .1141** **Bias = .0141** **Var(Est) = .0005** **Avg. T = .0006** **Coverage = .917**	**N = 990** **Est = .1184** **Bias = .0184** **Var(Est) = .0005** **Avg. T = .0007** **Coverage = .910**	**N = 1000** **Est = .1047** **Bias = .0047** **Var(Est) = .0002** **Avg. T = .0002** **Coverage = .947**	**N = 1000** **Est = .1065** **Bias = .0065** **Var(Est) = .0002** **Avg. T = .0002** **Coverage = .924**
DA estimate: 10 Imputed Datasets	**N = 1000** **Est = .1139** **Bias = .0139** **Var(Est) = .0004** **Avg. T = .0006** **Coverage = .928**	**N = 990** **Est = .1182** **Bias = .0182** **Var(Est) = .0005** **Avg. T = .0007** **Coverage = .911**	**N = 1000** **Est = .1052** **Bias = .0052** **Var(Est) = .0002** **Avg. T = .0002** **Coverage = .939**	**N = 1000** **Est = .1067** **Bias = .0067** **Var(Est) = .0002** **Avg. T = .0002** **Coverage = .932**
ML Estimate	Estimate = .0988 Var = .0005	Estimate = .0981 Var = .0005	Estimate = .1001 Var = .0004	Estimate = .0998 Var = .0002

Note: Bolded results indicate that bias is not significantly different from 0 with 95% confidence or bias is ≤ .02.

Table 7 (continued)

Simulation Results for Selected Parameters

TAU11: No Use at Time 1, No Use at Time 2 (τ = .7)

Weak rho parameters (.3 and .7):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	N = 76 Est = .4707 Bias = -.2293 Var(Est) = .0185 Avg. T = .0496 Coverage = .947	N = 38 Est = .4654 Bias = -.2346 Var(Est) = .0217 Avg. T = .0496 Coverage = .974	N = 676 Est = .6445 Bias = -.0555 Var(Est) = .0093 Avg. T = .0151 Coverage = .985	
DA estimate: 10 Imputed Datasets	N = 76 Est = .4653 Bias = -.2347 Var(Est) = .0167 Avg. T = .0434 Coverage = .961	N = 38 Est = .4695 Bias = -.2305 Var(Est) = .0199 Avg. T = .0443 Coverage = 1.00	N = 676 Est = .6456 Bias = -.0544 Var(Est) = .0081 Avg. T = .0142 Coverage = .984	
ML Estimate	Estimate = .6331 Var = .0490	Estimate = .6185 Var = .0601	Estimate = .6804 Var = .0164	

Strong rho parameters (.1 and .9):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	**N = 1000** **Est = .7093** **Bias = .0093** **Var(Est) = .0046** **Avg. T = .0068** **Coverage = .965**	**N = 990** **Est = .7050** **Bias = .0050** **Var(Est) = .0056** **Avg. T = .0076** **Coverage = .977**	**N = 1000** **Est = .7176** **Bias = .0176** **Var(Est) = .0023** **Avg. T = .0023** **Coverage = .945**	**N = 1000** **Est = .7182** **Bias = .0182** **Var(Est) = .0027** **Avg. T = .0028** **Coverage = .939**
DA estimate: 10 Imputed Datasets	**N = 1000** **Est = .7121** **Bias = .0121** **Var(Est) = .0043** **Avg. T = .0065** **Coverage = .983**	**N = 990** **Est = .7041** **Bias = .0041** **Var(Est) = .0052** **Avg. T = .0070** **Coverage = .980**	**N = 1000** **Est = .7165** **Bias = .0165** **Var(Est) = .0022** **Avg. T = .0021** **Coverage = .950**	**N = 1000** **Est = .7174** **Bias = .0174** **Var(Est) = .0025** **Avg. T = .0026** **Coverage = .950**
ML Estimate	Estimate = .6942 Var = .0055	Estimate = .6932 Var = .0067	Estimate = .7026 Var = .0018	Estimate = .7016 Var = .0021

Note: Bolded results indicate that bias is not significantly different from 0 with 95% confidence or bias is ≤ .02.

Table 7 (continued)

Simulation Results for Selected Parameters

TAU46: Alcohol+Cigarettes at Time 1, Alcohol+Cigarettes+Drunk at Time 2 (τ = .1)

Weak rho parameters (.3 and .7):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	N = 76 Est = .2234 Bias = .1234 Var(Est) = .0091 Avg. T = .0690 Coverage = 1.00	N = 38 Est = .2030 Bias = .1030 Var(Est) = .0108 Avg. T = .0641 Coverage = 1.00	N = 676 Est = .1959 Bias = .0959 Var(Est) = .0113 Avg. T = .0558 Coverage = .993	
DA estimate: 10 Imputed Datasets	N = 76 Est = .2236 Bias = .1236 Var(Est) = .0056 Avg. T = .0640 Coverage = 1.00	N = 38 Est = .2188 Bias = .1188 Var(Est) = .0069 Avg. T = .0668 Coverage = 1.00	N = 676 Est = .1952 Bias = .0952 Var(Est) = .0071 Avg. T = .0506 Coverage = 1.00	
ML Estimate	Estimate = .1736 Var = .0810	Estimate = .1854 Var = .0891	Estimate = .1476 Var = .0501	

Strong rho parameters (.1 and .9):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	N = 1000 Est = .1240 Bias = .0240 Var(Est) = .0069 Avg. T = .0178 Coverage = .988	N = 990 Est = .1371 Bias = .0371 Var(Est) = .0088 Avg. T = .0235 Coverage = .991	**N = 1000** **Est = .0958** **Bias = -.0042** **Var(Est) = .0030** **Avg. T = .0042** **Coverage = .972**	**N = 1000** **Est = .0968** **Bias = -.0033** **Var(Est) = .0039** **Avg. T = .0054** **Coverage = .966**
DA estimate: 10 Imputed Datasets	N = 1000 Est = .1264 Bias = .0264 Var(Est) = .0062 Avg. T = .0170 Coverage = .992	N = 990 Est = .1369 Bias = .0369 Var(Est) = .0074 Avg. T = .0222 Coverage = .996	**N = 1000** **Est = .0964** **Bias = -.0036** **Var(Est) = .0029** **Avg. T = .0038** **Coverage = .977**	**N = 1000** **Est = .0969** **Bias = -.0031** **Var(Est) = .0034** **Avg. T = .0051** **Coverage = .981**
ML Estimate	Estimate = .1081 Var = .0094	Estimate = .1067 Var = .0118	Estimate = .0975 Var = .0003	Estimate = .0987 Var = .0039

Note: Bolded results indicate that bias is not significantly different from 0 with 95% confidence or bias is $\leq$ .02.

Table 7 (continued)

Simulation Results for Selected Parameters

RHO15: 'No' to Marijuana Item, Given 'No Use' (ρ = .1 or .3)

Weak rho parameters (.3 and .7):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	N = 76 Est = .7214 Bias = .0214 Var(Est) = .0013 Avg. T = .0022 Coverage = .974	**N = 38** **Est = .7200** **Bias = .0200** **Var(Est) = .0017** **Avg. T = .0018** **Coverage = .974**	**N = 676** **Est = .7106** **Bias = .0106** **Var(Est) = .0005** **Avg. T = .0007** **Coverage = .957**	
DA estimate: 10 Imputed Datasets	N = 76 Est = .7206 Bias = .0206 Var(Est) = .0011 Avg. T = .0019 Coverage = .987	**N = 38** **Est = .7172** **Bias = .0172** **Var(Est) = .0016** **Avg. T = .0668** **Coverage = .974**	**N = 676** **Est = .7107** **Bias = .0107** **Var(Est) = .0005** **Avg. T = .0007** **Coverage = .969**	
ML Estimate	Estimate = .7108 Var = .0024	Estimate = .7178 Var = .0032	Estimate = .7033 Var = .0007	

Strong rho parameters (.1 and .9):

	300 Subjects		1000 Subjects	
	No Missing Data	Missing Data	No Missing Data	Missing Data
DA estimate: 5 Imputed Datasets	**N = 1000** **Est = .8927** **Bias = -.0073** **Var(Est) = .0004** **Avg. T = .0004** **Coverage = .937**	**N = 990** **Est = .8905** **Bias = -.0095** **Var(Est) = .0004** **Avg. T = .0005** **Coverage = .939**	**N = 1000** **Est = .8982** **Bias = -.0018** **Var(Est) = .0001** **Avg. T = .0042** **Coverage = .954**	**N = 1000** **Est = .8983** **Bias = -.0017** **Var(Est) = .0001** **Avg. T = .0002** **Coverage = .958**
DA estimate: 10 Imputed Datasets	**N = 1000** **Est = .8924** **Bias = -.0076** **Var(Est) = .0003** **Avg. T = .0004** **Coverage = .940**	**N = 990** **Est = .8905** **Bias = -.0095** **Var(Est) = .0004** **Avg. T = .0048** **Coverage = .939**	**N = 1000** **Est = .8985** **Bias = -.0015** **Var(Est) = .0001** **Avg. T = .0001** **Coverage = .960**	**N = 1000** **Est = .8984** **Bias = -.0016** **Var(Est) = .0001** **Avg. T = .0002** **Coverage = .965**
ML Estimate	Estimate = .9018 Var = .0004	Estimate = .9027 Var = .0005	Estimate = .9009 Var = .0001	Estimate = .9012 Var = .0001

Note: Bolded results indicate that bias is not significantly different from 0 with 95% confidence or bias is $\leq$ .02.

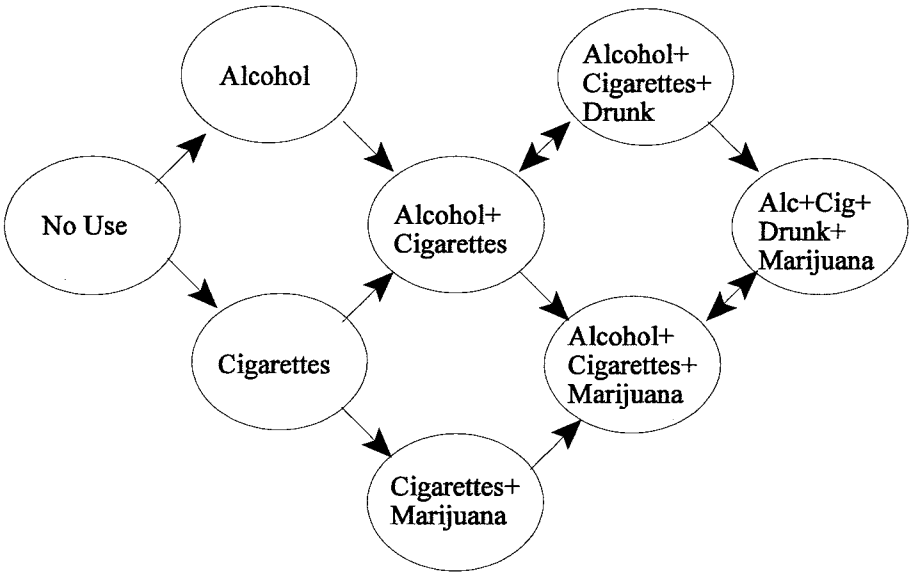

Figure1. The Hypothetical Model of Substance Use Onset

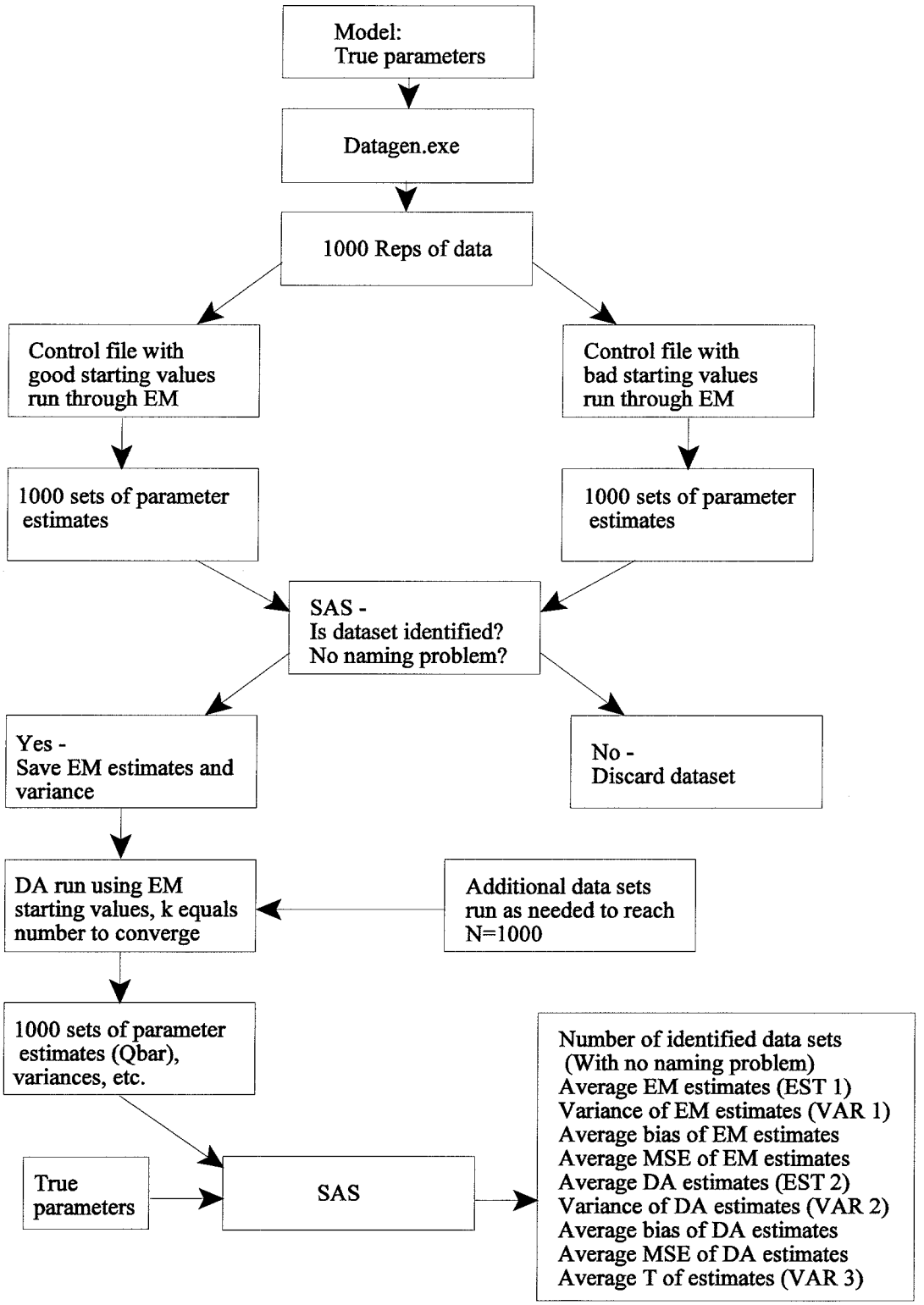

Figure 2. Flowchart of Simulation to Test LTA's Missing Data Routine