https://doi.org/10.1027/1015-5759/a000751

A Probit Multistate IRT Model With Latent Item Effect Variables for Graded Responses: Supplementary Material

October 13, 2022

Contents

A. Supplementary Material

A.1. A Probit Multistate IRT Model for Graded Responses

We define a multistate model for ordinal observables and call it the *probit multistate IRT model for graded responses* (PMG). However, additional assumptions for $\tilde{\eta}_t$ are necessary to estimate the parameters associated with $\tilde{\eta}_t$. One logical assumption to be made for the PMG model on the category level is that the (U_t, S_t)-conditional probabilities of answering in any of the categories $k = 1, \ldots, l$ or higher of the item i at time t are not equal. Also, an assumption made on the item level is that the (U_t, S_t)-conditional probabilities of answering in the category k (or higher) of any of the items $i = 1, \ldots, m$ at time t are not equal.

A.1.1. Assumptions of the PMG Model

For the definition of a probit multistate IRT model for graded responses, we introduce the following assumptions:

Essential Equivalence. The item-, time- and category-specific *threshold parameters* κ_{ikt} are introduced, which are real numbers. We assume that $\tilde{\tau}_{ikt}$ and $\tilde{\tau}_{i'k't}$ $\forall k, k' = 1, \ldots, l$, $\forall i, i' = 1, \ldots, m$, at time t only differ by a constant, which is a real number, when $(i, t) \neq (i', t')$ and $(i, t), (i', t') \in \{1, \ldots, m\} \times \{1, \ldots, n\}$. That is

$$\tilde{\tau}_{ikt} = \tilde{\eta}_t - \kappa_{ikt}, \;\; \forall\, i = 1, \ldots, m, \; \forall\, k = 1, \ldots, l, \; \forall\, t = 1, \ldots, n. \tag{1}$$

Every $\tilde{\tau}_{ikt}$ is therefore a translation of $\tilde{\eta}_t$ by a subtractive constant κ_{ikt}. If Equation 1 holds, the probit state variables $\tilde{\tau}_{ikt}$ are essentially equivalent with respect to $\tilde{\eta}_t$. There is always one threshold less for an item i at time t than there are possible response categories, which is why the first category index of a threshold parameter of an item i at a time t is $k = 1$.

Note that with this definition, every $\tilde{\tau}_{ikt}$ is perfectly determined by $\tilde{\eta}_t$ meaning that every $\tilde{\tau}_{ikt}$ can be computed from $\tilde{\eta}_t$ while in that equation, there is no error term. Therefore $\tilde{\eta}_t$ is not only the probit *reference* latent state variable but also the probit *common* latent state variable pertaining to time t.

For graded responses, we expect a hierarchical order of the threshold parameters within one item i at time t. For example, the value on $\tilde{\eta}_1$ of a random person-at-time-1 in a situation at time 1 is the probit transformed $(U_1 = u_1, S_1 = s_1)$-conditional probability of responding in category $k \geq 1$ to item $i = 1$ at time $t = 1$, that is, the value of $\tilde{\eta}_1$ is $\Phi^{-1}(P(\mathit{1}_{Y_{11} \geq 1} = 1 \mid U_1 = u_1, S_1 = s_1))$. An exemplary value of $\tilde{\eta}_1$ which is equal to 0.88 would transform to $P(\mathit{1}_{Y_{11} \geq 1} = 1 \mid U_1 = u_1, S_1 = s_1) = 0.81$. If there is, for example, a subtractive constant of $\kappa_{121} = 1.08$, then, for the category $k = 2$ of the same item $i = 1$ at time $t = 1$, there is a value of $\tilde{\tau}_{121} = \tilde{\eta}_1 - \kappa_{121}$ which is equal to -0.20. This value of $\tilde{\tau}_{121}$ translates into the $(U_1 = u_1, S_1 = s_1)$-conditional probability of responding in category $k \geq 2$ to item $i = 1$ at time $t = 1$, that is $P(\mathit{1}_{Y_{11} \geq 2} = 1 \mid U_1 = u_1, S_1 = s_1) = 0.42$. This example shows that a greater threshold parameter (here $\kappa_{121} > \kappa_{111} = 0$) yields smaller (U_t, S_t)-conditional probabilities of responding in a category k or higher. Within one item, the values of the threshold parameters κ_{ikt} are greater for higher category indexes k.

Equations 8 and 1 imply that $\kappa_{11t} = 0, \ \forall\, t = 1, \ldots, n$. Therefore, the threshold parameter associated with the reference category $k=1$ of the reference item $i=1$ has a value of 0. The threshold parameters of category $k=1$ of the other (non-reference) items κ_{i1t}, $\forall\, i=2, \ldots, m$, $\forall\, t=1, \ldots, n$, may be interpreted as the *time-specific item difficulty.* Note, that the other threshold parameters κ_{ikt}, $\forall\, k=2, \ldots, l$, are ignored for this concept of item difficulty.

If we assume that the lowest item category is always zero, the number of the highest item category l is always equal to the number of thresholds parameters per item. Because κ_{11t} is fixed to zero, for a number of items m and a number of measurement occasions n, the number of free threshold parameters is the product $m \cdot n \cdot (l-1) - m$.

Local independence. Just like in Section 3.3, we assume the (U_t, S_t)-conditional stochastic independence for the PMG model (see Equation 15). According to this assumption, given U_t and S_t, each response variable Y_{it} is independent of the other response variables $Y_{i't'}$, of past and future person variables, independent of past and future situation variables, and independent of all subsets of all these variables.

Figure 1 shows a path diagram of the PMG model. This path diagram displays the decomposition of $1_{Y_{it} \geq k}$ into the conditional probability τ_{ikt} and a measurement error variable ϵ_{ikt}. The conditional probability τ_{ikt} is transformed by the link function into its probit $\tilde{\tau}_{ikt}$. This is indicated by the curve connecting τ_{ikt} and $\tilde{\tau}_{ikt}$. The probit state variable τ_{ikt} is perfectly determined by the probit reference latent state variable $\tilde{\eta}_t$ which means that there is no residual variable in the decomposition of $\tilde{\tau}_{ikt}$ into the difference between $\tilde{\eta}_t$ and κ_{ikt} (see Equation 1).

A.2. Application

A.2.1. Analysis of the (Unconstrained) PIEG-1 Model: Real Data Analysis

In this section, we interpret the results of the unconstrained PIEG model (see Section A.2.3) and we compare it with the PMG model to investigate if latent item effect variables should be included in the model. With the PMG as well as with the PIEG model, we estimate the variances, expectations, and covariances of one reference latent state variable ($\tilde{\eta}_t$) for each time point of measurement. As the reference item, we used the item *satisfied.* In the PIEG model, in contrast to the PMG model, we also estimate the parameters of one latent item effect variable ($\tilde{\beta}_i$) for each item, except for the reference item, resulting in a total of three reference latent state variables and three latent item effect variables. Note that they are defined on the probit level. As the items include five categories, 36 free threshold parameters (κ_{ikt}) are estimated with the unconstrained PIEG model because the first threshold parameter of each item at each time point of

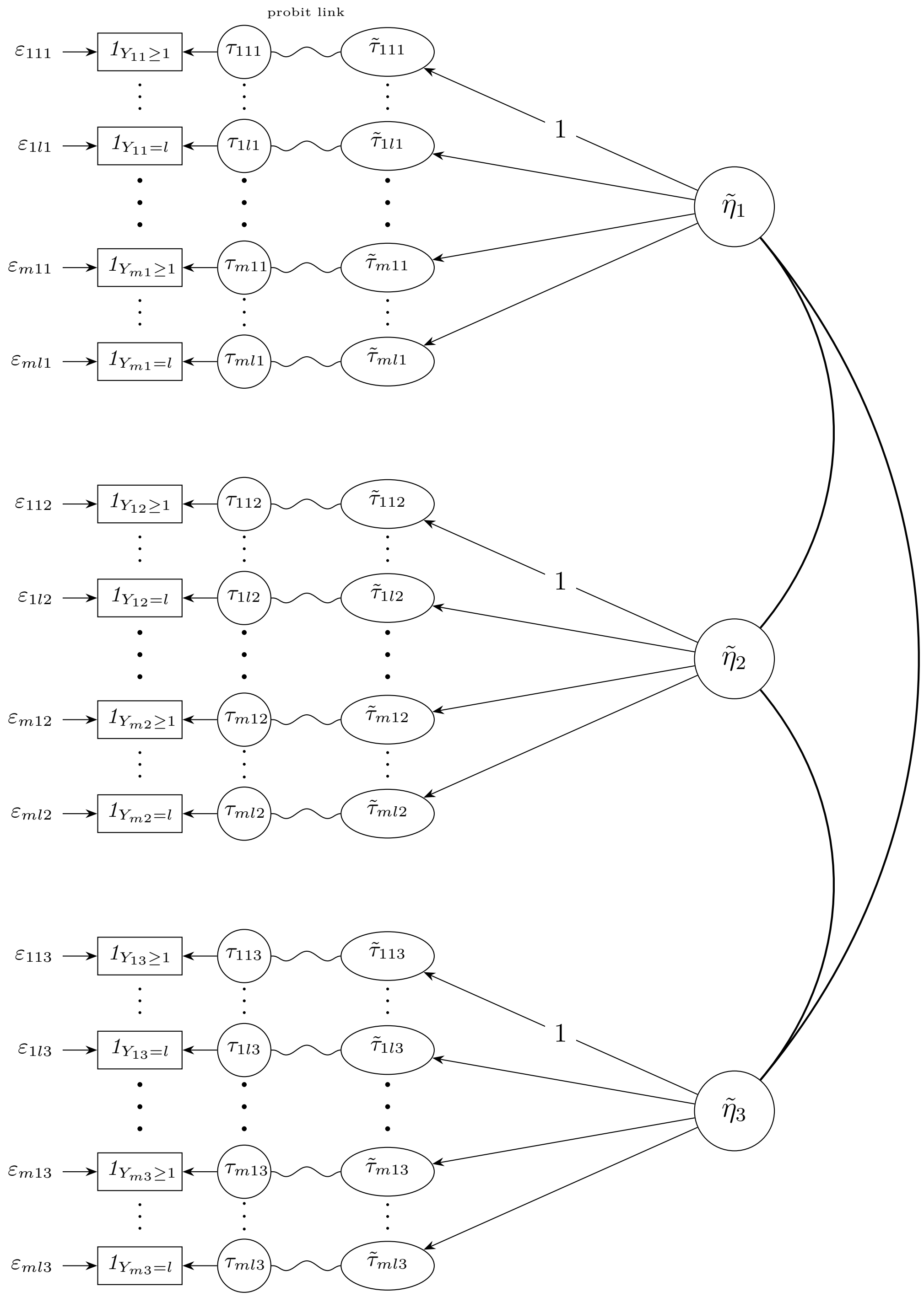

Figure 1. Path diagram of the probit multistate IRT model for graded responses with $m \cdot l \cdot 3$ indicator variables $1_{Y_{it}\geq k}$ measuring three probit reference latent state variables $\tilde{\eta}_t$.

measurement is fixed to zero ($\kappa_{i1t}=0$). In the PMG model, 45 free threshold parameters are estimated as only the first threshold parameter at each time point is fixed to zero ($\kappa_{11t}=0$).

Results The PMG model does not fit the data ($\chi^2(60)=458.621, p=.000$, RMSEA = 0.116). The PIEG model, on the other hand, fits the data very well ($\chi^2(51)=61.516, p=.149$, RMSEA = .020). The model fit is significantly better than the model fit of the PMG model ($LR=254.077, df=9, p<.000$). According to these results, including latent item effect variables considerably improves the model fit. This is an indication that the items do not only differ with respect to constant difficulty. Instead, each non-reference item has a latent item effect variable. One of the interpretations of the values of this latent item effect variable is that they are person-specific item difficulties. That is the case, although the items were constructed to measure the latent state variable equally well for all persons. Also, every item seems to have specific characteristics deviating from the latent state.

We now interpret the certain parameter estimates of the (unconstrained) PIEG-1 model. First, we consider the estimated parameters of the reference latent state variables that represent current well-being at all occasions of measurement. The reference latent state variable is defined as the probit transformed (U_t, S_t)-conditional probability of answering in the reference category ($k=1$) or higher of the reference item *satisfied* ($i=1$). Expectations, variances and covariances of the three reference latent state variables significantly differ from zero.

Furthermore, we examine the estimated parameters of the latent item effect variables. The latent item effect variable is defined as the difference between the probit transformed latent state variable of a non-reference item i and the probit transformed latent state variable of the reference item $i=1$ (see Equation 13). A positive value of a person on the item effect variable indicates that it is less probable for that person to respond in the reference category ($k=1$) or higher of the reference item ($i=1$) than in the reference category or higher of the non-reference item i.

The estimated variances of the latent item effect variables all significantly differ from zero (see Table 1). This supports the hypothesis that there are person-specific effects of assessing current well-being with a non-reference item instead of the reference item. Also, the estimated means significantly differ from zero (see Table 1). This indicates that there are non-zero average item-specific effects. The expectation of the probit state variable of the reference category (see Equation 10) is $E(\tilde{\tau}_{it}) = E(\tilde{\beta}_i) + E(\tilde{\eta}_t)$. Analogously, this applies to the estimate of the average of the probit state variable of the reference category. The difference between the estimates of $E(\tilde{\tau}_{1t})=E(\tilde{\eta}_t)$ and $E(\tilde{\tau}_{it})$ for the non-reference item i at time t is the estimate of $E(\tilde{\beta}_i)$. These estimates differ between items but are identical for all time points and can be interpreted as estimates of the average of the individual item difficulties (see Section 3.2). In our case, the means of the item effect variables of the items *unsatisfied* ($i=2$) and *unhappy* ($i=4$) are positive. These estimated positive expectations mean that it is easier for a randomly sampled person to respond to these items, as compared to the positively worded reference item, in category $k = 1$ or higher.

Furthermore, our application shows that the correlations between the item effect variables $\tilde{\beta}_i$ and the reference latent state variables $\tilde{\eta}_t$ are not significantly different from zero (see Table 1). This is an indication that variation of the person-specific item effects does not go along with the variation of general current well-being assessed with the reference item.

A.2.2. Monte Carlo Simulation of the (Unconstrained) PIEG-1 Model

In Monte Carlo simulations, many samples of data are generated given a specified model. The samples are then analyzed and the results are aggregated across all generated data sets. Monte Carlo simulations are conducted to test the quality of parameter estimates if certain conditions, such as the sample size, are varied. This allows us to investigate how sample sizes affect the estimation of the parameters, standard errors, or model-fit indices.

Several data sets, which are compatible with the PIEG-1 model, were generated. The generated data sets consist of five-category response variables for 4 items at 3 time points of measurement (12 response variables). We used M*plus* (Muthén & Muthén, 1998) for data generation. The model structure of the PIEG-1 model (see Supplementary Material A.2) together with the true parameters make up the M*plus* input.

We conducted six different simulations for different sample sizes N_{obs} ranging from 250 to $10,000$ for the PIEG-1 model, which has no threshold parameter constraints. We requested $N_{rep} = 1,000$ replications.

The generated data sets were analyzed individually using the WLSMV estimator. The resulting estimates were aggregated for each sample size across the 1,000 generated data sets. For example, for all parameter estimates and corresponding standard error estimates within one sample size, the mean and the standard deviation were calculated across all replications. Because of identification and convergence problems, for smaller sample sizes fewer replications were successfully produced in the Monte Carlo simulation.

In order to evaluate the quality of the parameter estimation, we checked how well the true parameters can be recovered in the analysis. For different sample sizes, we assess the statistical performance of each parameter of the latent variables as well as the model fit indices of the PIEG-1 model (for 4 items at 3 measurement occasions). We evaluate the statistical performance on the basis of different indices suggested by B. Muthén and Muthén (2002): the parameter estimate bias (*peb*), the standard error bias (*seb*), coverage, and the distributional properties of the χ^2 statistics. For the PIEG model to be preferable over the PMG model, the variances of $\tilde{\beta}_i$ must be greater than 0. For the hypotheses that $Var(\tilde{\beta}_i)$ are greater than 0, we also check the statistical power.

The parameter estimate bias *peb* is a standardized measure to describe the discrepancies between a sample estimate and the corresponding true parameter. It is defined as the relative deviation of the average of the sample estimates of the parameter from the true parameter:

$$peb = \frac{m_p - e_p}{e_p}, \tag{2}$$

where m_p is the mean of the parameter estimates across replications, and e_p is the true parameter. The parameter estimate bias *peb* should be understood as an index for relative bias. For a true parameter of 0, the *peb* is not defined. For true parameters close to 0, the absolute bias $m_p - e_p$ may be of interest rather than the *peb*.

The standard error bias *seb* is a standardized measure of the accuracy of standard errors, defined as

$$seb = \frac{m_{se} - SD_p}{SD_p}, \tag{3}$$

where m_{se} is the mean of the estimates of the standard error across replications. The standard deviation from the parameter estimates across replications SD_p approximates the true standard error when the number of replications is large. Therefore the *seb* is the relative deviation of the average standard error estimate and the proxy for the corresponding true standard error.

Coverage is the proportion of replications for which the 95% confidence interval contains the true parameter value. It indicates the accuracy of confidence interval estimation. Over- or underestimation of standard errors leads to a coverage deviating from .95.

According to B. Muthén and Muthén (2002), parameter and standard error estimates have a good statistical performance if *peb* and *seb* values are smaller than .10 and coverage remains between .91 and .98. As to the model test, rejection rates should be close to .05 with respect to the χ^2 distribution. For the PIEG-1 model, we used the χ^2 distribution with 51 degrees of freedom (see Section A.2.1). For the estimates of the variances of the latent item effect variables $\tilde{\beta}_i$, we also assessed power as the proportion of replications for which the null hypothesis, that the true parameter is equal to zero, was rejected. Of course, power was only assessed if the true parameters corresponding to the estimates were not 0. According to B. Muthén and Muthén (2002), power estimates over .80 indicate good statistical performance of the null hypothesis test. Also, for the parameters for which power was assessed, *seb* values smaller than .05 indicate satisfactory statistical performance of the parameter estimates.

We used the rounded parameter estimates of the preceding analysis of the PIEG-2 model (see Section A.2.3) as true parameters. Six different simulations for different sample sizes N_{obs} ranging from 250 to $10,000$ with $N_{rep} = 1000$ replications for each sample size were conducted. We evaluated the parameter estimate bias (*peb*), the standard error bias (*seb*), and the coverage of all latent variable means, variances, and covariances as well as the distributional properties of χ^2 statistics. For the hypotheses that $Var(\beta_i)$ is not equal to 0, we also check the statistical power.

In Table 2 and 3, we refer to power as the proportion of replications for which the null hypothesis, that the true parameter is equal to zero, is rejected at the .05 level (when a two-tailed z-test[1] with a critical value of 1.96 is used). If the number of replications is

[1] For some parameter constellations, however, the standard normal distribution may not be an adequate

large and the true parameter in question is different from zero, this value estimates the probability of rejecting the null hypothesis.

Results For all parameters and sample sizes, there were no parameter estimates with a coverage smaller than .91 or higher than .98, nor *seb* values greater than .10 (see Table 3). For all $Var(\tilde{\beta}_i)$, the *seb* values were smaller than 0.05. These results suggest that there is satisfactory standard error estimation for the PIEG-1 model (for four 5-category items at three time points) even if sample sizes are as small as $n = 250$. We consider these results to be satisfactory.

Only for very small true parameters, *peb* values were greater than .10. Here, the absolute bias $m_p - e_p$ is of interest. In our simulation, absolute values of the true parameters for all $Cov(\tilde{\beta}_i, \tilde{\eta}_t)$ had a range of 0.01 to 0.16. The estimates for these parameters showed an absolute bias of < 0.01 for all sample sizes. This means that even the parameter estimates of the simulated samples with a sample size of $n = 250$ were, on average, not exceeding a difference of 0.01 to the true parameter even if the true parameter was very small. For example, the parameter $Cov(\tilde{\beta}_4, \tilde{\eta}_2)$ has a true parameter of 0.01. In our simulation, the estimation of $Cov(\tilde{\beta}_4, \tilde{\eta}_2)$ showed an absolute bias of -0.01 for the sample size of $n = 250$. This indicates satisfactory parameter estimation of the PIEG-1 model, even for small sample sizes and small parameters.

The true parameter, the average of the parameter estimate and of the standard error estimate across replications, the parameter estimate bias, the standard error bias as well as coverage and power for every latent variable mean and variance of all the PIEG-2 model Monte Carlo simulations are shown in Table 2. Results for the covariance estimates are shown in Table 3.

Table 2 about here.
Table 3 about here.

To investigate the distributional properties of the χ^2 statistics for the model fit test, rejection rates, which were based on a Type i error α-level of .05 of the model test, were analyzed. They indicated that, for small sample sizes, the PIEG-1 model tends to slightly overestimate the goodness of fit. Only 4% of the simulated χ^2 values exceeded the cutoff value of the $95th$ percentile of the χ^2 distribution with 51 degrees of freedom. This may suggest that the power of the χ^2 goodness-of-fit model test may suffer if the sample size is smaller than $n = 1000$.

The averages of the χ^2 value estimates across replications, and rejection rates are depicted in Table 1.

Table 1 about here.

test function.

A.2.3. Monte Carlo Simulation of the PIEG-2 Model

The PIEG-2 model, as the more restrictive model compared to the PIEG-1 model, fit the data well in real data application. Now we study the quality of the parameter estimates of the PIEG-2 model via Monte Carlo simulations. In Supplementary Material A.2.2, the statistical performance of the unrestricted PIEG-1 model is investigated. The results of the simulation show that the model estimation performs relatively well even for small sample sizes.

Results The true parameter, the average of the parameter estimate and of the standard error estimate across replications, the parameter estimate bias, the standard error bias as well as coverage and power for every latent variable mean and variance of all the PIEG-2 model Monte Carlo simulations are shown in Table 4. Results for the covariance estimates are shown in Table 5.

Table 4 about here.

Table 5 about here.

Coverage and standard error biases do not exceed the cutoff criteria proposed by B. Muthén and Muthén (2002) for any of the parameter estimates for any of the sample sizes. These results indicate a good statistical performance of the standard error estimates so that the standard errors are probably not highly over- nor underestimated with the PIEG-2 model. Also, the parameter estimate biases *peb* only exceeded 0.05 for the covariance estimates which have very small true parameters. However, the absolute biases for the estimates of all covariances are < 0.01 for all sample sizes. This means that the Monte Carlo simulation results suggest good statistical performance of all the parameter estimates of the PIEG-2 model. We therefore assume that the threshold parameter constraint $\kappa_{ikt} = \kappa_{ik1} \;\forall\, k = 1, \ldots, l,\;\; \forall\, i = 1, \ldots, m,\;\; \forall\, t = 1, \ldots, n$, does not largely affect the statistical performance of the parameter estimates.

The PIEG-2 model simulation produces empirical χ^2 distributions that are slightly shifted to the left if the sample sizes are small. The rejection rates, shown in Table 1, are close to 4% for sample sizes smaller than 750. This may suggest that the power of the χ^2 goodness-of-fit test for the PIEG-2 model may diminish if the sample size is smaller than $n = 750$.

A.2.4. Interpretation of the Monte Carlo Simulations

Overall, the Monte Carlo simulations showed that the WLSMV estimator provided very stable results and that this method seems to be adequate for the estimation of the PIEG-1 and the PIEG-2 model with 12 response variables. The quality of the parameter estimates, standard errors and model fit indices associated with the PIEG-1 and PIEG-2 model using the WLSMV estimator (B. O. Muthén et al., 2015) is generally very good. For both the PIEG-1 and the PIEG-2 model, the standard error biases are negligibly small for all parameter estimates for all sample sizes of the simulated samples. Also, for both models, for all parameters, and for all sample sizes there are no parameter estimates with a coverage smaller than .91 or higher than .98. These results indicate a good statistical performance of the standard error estimates of both the PIEG-1 and the PIEG-2 model even if sample sizes are small. In order to illustrate how the sample size affects standard error estimation, we conducted additional Monte Carlo simulations for the PIEG-1 as well as for the PIEG-2 model for sample sizes of $n = 100$. Figure 2 shows that the mean standard error estimates for $Var(\tilde{\beta}_2)$ across the simulated samples decrease with increasing sample sizes for the PIEG-1. The decline of these mean standard error estimates seems particularly pronounced for sample sizes under $n = 1000$. Figure 3 shows all mean standard error estimates for the expectations and variances of the latent variables of the PIEG-1 model.

Figure 2 about here.
Figure 3 about here.

The point estimates of the parameters also showed a satisfying quality for all sample sizes for both models in the Monte Carlo simulation. The true covariances $Cov(\tilde{\eta}_t, \tilde{\beta}_i)$ were very close to 0. For the estimates of these parameters, the absolute biases $m_p - e_p$ were < 0.01, which we considered to be satisfactory.

References

Muthén, B. O., Muthén, L., & Asparouhov, T. (2015). Estimator choices with categorical outcomes. *Mplus Technical Appendix*, *8*.

Muthén, B., & Muthén, L. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*(4), 599–620.

Muthén & Muthén. (1998). *Mplus: The comprehensive modeling program for applied researchers: User's guide*. Muthén & Muthén.

A.1. Tables

Table 1

Monte Carlo simulation results for χ^2 values of the PIEG-1 model and the PIEG-2 model.

Model	N_{obs}	N_{repl}	df	Mean χ^2	Rej. Rate
PIEG-1	250	996	51	50.924	0.046
PIEG-1	500	1000	51	50.832	0.042
PIEG-1	750	1000	51	51.077	0.044
PIEG-1	1000	1000	51	51.379	0.058
PIEG-1	5000	1000	51	51.085	0.054
PIEG-1	10000	1000	51	50.583	0.050
PIEG-2	250	997	75	75.249	0.039
PIEG-2	500	1000	75	74.844	0.034
PIEG-2	750	1000	75	75.277	0.048
PIEG-2	1000	1000	75	74.803	0.052
PIEG-2	5000	1000	75	74.915	0.056
PIEG-2	10000	1000	75	74.588	0.052

Table 2

Results of the Monte Carlo simulation for means and variances of the latent variables in the PIEG-1 model.

		Mean						Variance					
		250	500	750	1000	5000	10000	250	500	750	1000	5000	10000
$\tilde{\eta}_1$	True Param. (e_p)	3.54						2.85					
	Mean (m_p)	3.605	3.571	3.561	3.558	3.541	3.541	2.993	2.921	2.898	2.885	2.857	2.853
	Bias (peb)	0.018	0.009	0.006	0.005	0.000	0.000	0.050	0.025	0.017	0.012	0.003	0.001
	S.E. Mean (m_{se})	0.258	0.179	0.145	0.126	0.056	0.039	0.468	0.322	0.261	0.225	0.100	0.071
	S.E. Bias (seb)	0.013	0.011	0.006	0.001	0.014	0.029	0.006	0.044	0.030	0.022	0.009	0.002
	Coverage	0.946	0.954	0.950	0.958	0.957	0.951	0.948	0.942	0.950	0.948	0.943	0.947
	Power	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\eta}_2$	True Param. (e_p)	3.95						3.86					
	Mean (m_p)	4.030	3.990	3.977	3.974	3.955	3.952	4.045	3.938	3.906	3.893	3.867	3.861
	Bias (peb)	0.020	0.010	0.007	0.006	0.001	0.001	0.048	0.020	0.012	0.009	0.002	0.000
	S.E. Mean (m_{se})	0.283	0.197	0.160	0.138	0.061	0.043	0.620	0.426	0.345	0.298	0.133	0.094
	S.E. Bias (seb)	0.009	0.034	0.042	0.006	0.036	0.006	0.058	0.025	0.008	0.025	0.011	0.016
	Coverage	0.944	0.962	0.960	0.947	0.958	0.949	0.945	0.943	0.953	0.949	0.947	0.949
	Power	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\eta}_3$	True Param. (e_p)	3.77						3.60					
	Mean (m_p)	3.851	3.817	3.799	3.795	3.775	3.772	3.774	3.673	3.644	3.635	3.606	3.606
	Bias (peb)	0.022	0.012	0.008	0.007	0.001	0.001	0.048	0.020	0.012	0.010	0.002	0.002
	S.E. Mean (m_{se})	0.273	0.190	0.154	0.133	0.059	0.042	0.576	0.398	0.322	0.278	0.123	0.087
	S.E. Bias (seb)	0.014	0.034	0.033	0.019	0.007	0.013	0.022	0.001	0.006	0.061	0.025	0.023
	Coverage	0.946	0.943	0.943	0.944	0.950	0.956	0.948	0.949	0.948	0.965	0.958	0.945
	Power	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\beta}_2$	True Param. (e_p)	0.57						1.10					
	Mean (m_p)	0.597	0.578	0.577	0.572	0.570	0.571	1.167	1.130	1.120	1.116	1.105	1.103
	Bias (peb)	0.048	0.014	0.012	0.004	0.001	0.002	0.061	0.027	0.018	0.015	0.005	0.002
	S.E. Mean (m_{se})	0.268	0.186	0.151	0.131	0.058	0.041	0.269	0.185	0.150	0.129	0.057	0.040
	S.E. Bias (seb)	0.026	0.010	0.004	0.001	0.017	0.047	0.023	0.012	0.016	0.001	0.001	0.018
	Coverage	0.941	0.943	0.953	0.949	0.941	0.966	0.955	0.954	0.959	0.952	0.951	0.945
	Power	0.618	0.882	0.972	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\beta}_3$	True Param. (e_p)	-0.37						1.09					
	Mean (m_p)	-0.368	-0.370	-0.365	-0.368	-0.369	-0.369	1.165	1.123	1.115	1.111	1.095	1.093
	Bias (peb)	-0.006	0.000	-0.012	-0.006	-0.003	-0.002	0.069	0.030	0.022	0.019	0.005	0.002
	S.E. Mean (m_{se})	0.229	0.160	0.130	0.112	0.050	0.035	0.251	0.172	0.139	0.120	0.053	0.037
	S.E. Bias (seb)	0.010	0.019	0.012	0.028	0.072	0.052	0.006	0.009	0.001	0.010	0.028	0.050
	Coverage	0.954	0.954	0.951	0.951	0.963	0.962	0.949	0.945	0.949	0.961	0.949	0.956
	Power	0.359	0.656	0.810	0.916	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\beta}_4$	True Param. (e_p)	0.84						1.34					
	Mean (m_p)	0.872	0.850	0.850	0.848	0.842	0.841	1.425	1.382	1.369	1.365	1.346	1.343
	Bias (peb)	0.038	0.012	0.012	0.009	0.002	0.001	0.063	0.031	0.022	0.019	0.004	0.002
	S.E. Mean (m_{se})	0.294	0.204	0.166	0.143	0.063	0.045	0.317	0.219	0.178	0.154	0.068	0.048
	S.E. Bias (seb)	0.017	0.019	0.036	0.001	0.026	0.036	0.035	0.003	0.008	0.021	0.058	0.018
	Coverage	0.956	0.949	0.959	0.959	0.953	0.953	0.950	0.947	0.953	0.942	0.961	0.950
	Power	0.862	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3

Results of the Monte Carlo simulation for covariances of the latent variables in the PIEG-1 model.

N_{obs}	Parameter	$Cov(\tilde{\eta}_1, \tilde{\eta}_2)$	$Cov(\tilde{\eta}_1, \tilde{\eta}_3)$	$Cov(\tilde{\eta}_2, \tilde{\eta}_3)$	$Cov(\tilde{\beta}_2, \tilde{\beta}_3)$	$Cov(\tilde{\beta}_2, \tilde{\beta}_4)$	$Cov(\tilde{\beta}_3, \tilde{\beta}_4)$	$Cov(\tilde{\beta}_2, \tilde{\eta}_1)$	$Cov(\tilde{\beta}_2, \tilde{\eta}_2)$	$Cov(\tilde{\beta}_2, \tilde{\eta}_3)$	$Cov(\tilde{\beta}_3, \tilde{\eta}_1)$	$Cov(\tilde{\beta}_3, \tilde{\eta}_2)$	$Cov(\tilde{\beta}_3, \tilde{\eta}_3)$	$Cov(\tilde{\beta}_4, \tilde{\eta}_1)$	$Cov(\tilde{\beta}_4, \tilde{\eta}_2)$	$Cov(\tilde{\beta}_4, \tilde{\eta}_3)$
	True Param. (e_p)	**1.00**	**0.77**	**1.22**	**0.15**	**1.15**	**0.3**	**0.12**	**0.12**	**0.04**	**-0.16**	**-0.12**	**-0.16**	**0.03**	**0.01**	**0.03**
250	Mean (m_p)	1.038	0.803	1.279	0.171	1.214	0.325	0.117	0.114	0.031	−0.174	−0.126	−0.169	0.035	0.000	0.016
	Bias (*peb*)	0.038	0.043	0.048	0.139	0.055	0.082	−0.025	−0.049	−0.227	0.085	0.048	0.055	0.164	−0.976	−0.455
	S.E. Mean (m_{se})	0.314	0.297	0.350	0.159	0.230	0.176	0.221	0.249	0.243	0.214	0.241	0.234	0.241	0.272	0.264
	S.E. Bias (*seb*)	0.032	0.002	0.012	0.030	0.073	0.054	0.025	0.022	0.057	0.017	0.032	0.045	0.027	0.027	0.030
	Coverage	0.946	0.950	0.958	0.949	0.933	0.940	0.952	0.949	0.939	0.954	0.951	0.939	0.949	0.944	0.938
	Power	0.946	0.808	0.990	0.167	1.000	0.439	0.090	0.086	0.063	0.097	0.073	0.106	0.050	0.057	0.061
500	Mean (m_p)	1.020	0.784	1.244	0.156	1.181	0.310	0.114	0.115	0.039	−0.167	−0.126	−0.160	0.029	0.003	0.025
	Bias (*peb*)	0.020	0.018	0.019	0.042	0.027	0.032	−0.047	−0.040	−0.037	0.045	0.048	0.002	−0.026	−0.736	−0.167
	S.E. Mean (m_{se})	0.220	0.208	0.245	0.110	0.160	0.122	0.154	0.174	0.168	0.149	0.167	0.163	0.168	0.190	0.184
	S.E. Bias (*seb*)	0.055	0.033	0.014	0.014	0.003	0.058	0.016	0.006	0.039	0.030	0.006	0.021	0.029	0.038	0.028
	Coverage	0.942	0.944	0.944	0.943	0.945	0.924	0.953	0.947	0.950	0.949	0.949	0.947	0.951	0.949	0.953
	Power	1.000	0.978	1.000	0.270	1.000	0.732	0.131	0.106	0.067	0.185	0.101	0.150	0.054	0.051	0.050
750	Mean (m_p)	1.013	0.773	1.232	0.156	1.172	0.308	0.118	0.117	0.041	−0.164	−0.120	−0.158	0.030	0.004	0.027
	Bias (*peb*)	0.013	0.004	0.010	0.039	0.019	0.025	−0.019	−0.025	0.031	0.024	−0.002	−0.013	−0.012	−0.555	−0.107
	S.E. Mean (m_{se})	0.179	0.169	0.199	0.090	0.130	0.099	0.125	0.141	0.137	0.121	0.136	0.133	0.137	0.154	0.149
	S.E. Bias (*seb*)	0.015	0.009	0.017	0.023	0.014	0.036	0.019	0.040	0.006	0.012	0.022	0.001	0.027	0.023	0.011
	Coverage	0.955	0.947	0.959	0.941	0.948	0.949	0.947	0.965	0.950	0.947	0.934	0.947	0.940	0.961	0.947
	Power	1.000	1.000	1.000	0.405	1.000	0.890	0.163	0.136	0.061	0.258	0.129	0.200	0.065	0.038	0.060
1000	Mean (m_p)	1.009	0.771	1.228	0.154	1.168	0.305	0.119	0.119	0.043	−0.166	−0.120	−0.160	0.029	0.011	0.031
	Bias (*peb*)	0.009	0.002	0.007	0.025	0.016	0.016	−0.011	−0.009	0.081	0.037	−0.003	−0.001	−0.028	0.118	0.042
	S.E. Mean (m_{se})	0.155	0.147	0.172	0.077	0.112	0.086	0.108	0.122	0.118	0.105	0.118	0.115	0.119	0.133	0.129
	S.E. Bias (*seb*)	0.025	0.010	0.000	0.017	0.022	0.043	0.002	0.008	0.018	0.020	0.024	0.016	0.017	0.026	0.021
	Coverage	0.947	0.942	0.944	0.949	0.940	0.938	0.949	0.956	0.954	0.945	0.944	0.943	0.945	0.949	0.951
	Power	1.000	1.000	1.000	0.510	1.000	0.953	0.202	0.177	0.071	0.329	0.169	0.275	0.064	0.053	0.060
5000	Mean (m_p)	1.005	0.772	1.221	0.150	1.155	0.300	0.121	0.119	0.043	−0.162	−0.121	−0.161	0.028	0.008	0.031
	Bias (*peb*)	0.005	0.002	0.001	0.001	0.004	0.000	0.007	−0.007	0.073	0.011	0.005	0.008	−0.078	−0.183	0.022
	S.E. Mean (m_{se})	0.069	0.065	0.077	0.034	0.050	0.038	0.048	0.054	0.053	0.046	0.052	0.051	0.053	0.059	0.058
	S.E. Bias (*seb*)	0.019	0.007	0.005	0.010	0.011	0.009	0.014	0.009	0.003	0.024	0.022	0.021	0.010	0.012	0.027
	Coverage	0.950	0.947	0.944	0.947	0.956	0.958	0.958	0.946	0.956	0.955	0.945	0.953	0.947	0.949	0.956
	Power	1.000	1.000	1.000	0.994	1.000	1.000	0.714	0.601	0.134	0.948	0.626	0.897	0.082	0.052	0.078
10000	Mean (m_p)	1.001	0.770	1.222	0.149	1.152	0.299	0.121	0.121	0.042	−0.159	−0.119	−0.160	0.030	0.010	0.030
	Bias (*peb*)	0.001	0.000	0.001	−0.003	0.002	−0.003	0.011	0.007	0.053	−0.005	−0.007	−0.001	−0.016	0.031	0.007
	S.E. Mean (m_{se})	0.049	0.046	0.054	0.024	0.035	0.027	0.034	0.038	0.037	0.033	0.037	0.036	0.037	0.042	0.041
	S.E. Bias (*seb*)	0.017	0.001	0.022	0.001	0.006	0.013	0.013	0.023	0.033	0.019	0.028	0.018	0.034	0.023	0.012
	Coverage	0.952	0.946	0.947	0.942	0.950	0.950	0.950	0.943	0.934	0.955	0.944	0.944	0.967	0.946	0.952
	Power	1.000	1.000	1.000	1.000	1.000	1.000	0.948	0.876	0.207	0.999	0.887	0.995	0.107	0.064	0.112

Table 4

Results of the Monte Carlo simulation for means and variances of the latent variables in the PIEG-2 model

		Mean						Variance					
	Index	250	500	750	1000	5000	10000	250	500	750	1000	5000	10000
$\tilde{\eta}_1$	True Param. (e_p)	3.63						3.11					
	Mean (m_p)	3.698	3.660	3.649	3.644	3.632	3.631	3.251	3.186	3.162	3.141	3.117	3.117
	Bias (peb)	0.019	0.008	0.005	0.004	0.000	0.000	0.045	0.024	0.017	0.010	0.002	0.002
	S.E. Mean (m_{se})	0.228	0.158	0.128	0.111	0.049	0.035	0.466	0.322	0.261	0.225	0.100	0.070
	S.E. Bias (seb)	0.018	0.010	0.017	0.002	0.011	0.009	0.019	0.034	0.040	0.020	0.017	0.030
	Coverage	0.954	0.955	0.955	0.954	0.951	0.952	0.946	0.953	0.950	0.944	0.947	0.960
	Power	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\eta}_2$	True Param. (e_p)	3.87						3.65					
	Mean (m_p)	3.937	3.905	3.891	3.886	3.873	3.871	3.812	3.731	3.703	3.683	3.660	3.655
	Bias (peb)	0.017	0.009	0.006	0.004	0.001	0.000	0.044	0.022	0.014	0.009	0.003	0.001
	S.E. Mean (m_{se})	0.237	0.164	0.133	0.115	0.051	0.036	0.544	0.376	0.304	0.262	0.116	0.082
	S.E. Bias (seb)	0.008	0.019	0.014	0.002	0.028	0.005	0.005	0.022	0.026	0.003	0.019	0.010
	Coverage	0.944	0.936	0.951	0.952	0.940	0.947	0.951	0.955	0.955	0.946	0.940	0.947
	Power	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\eta}_3$	True Param. (e_p)	3.75						3.52					
	Mean (m_p)	3.819	3.785	3.772	3.767	3.752	3.750	3.661	3.587	3.566	3.549	3.525	3.523
	Bias (peb)	0.018	0.009	0.006	0.004	0.000	0.000	0.040	0.019	0.013	0.008	0.001	0.001
	S.E. Mean (m_{se})	0.234	0.162	0.131	0.114	0.050	0.036	0.521	0.360	0.292	0.251	0.111	0.079
	S.E. Bias (seb)	0.019	0.016	0.006	0.000	0.004	0.014	0.021	0.011	0.003	0.002	0.010	0.014
	Coverage	0.953	0.950	0.945	0.941	0.959	0.953	0.960	0.944	0.954	0.952	0.949	0.948
	Power	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\tilde{\beta}_2$	True Param. (e_p)	0.56						1.1					
	Mean (m_p)	0.570	0.566	0.564	0.563	0.563	0.562	1.162	1.131	1.119	1.113	1.103	1.102
	Bias (peb)	0.018	0.011	0.007	0.006	0.005	0.003	0.056	0.028	0.018	0.012	0.003	0.002
	S.E. Mean (m_{se})	0.265	0.185	0.151	0.130	0.058	0.041	0.266	0.184	0.149	0.128	0.057	0.040
	S.E. Bias (seb)	0.023	0.057	0.020	0.036	0.019	0.008	0.028	0.026	0.025	0.011	0.046	0.042
	Coverage	0.951	0.948	0.948	0.941	0.956	0.948	0.954	0.949	0.947	0.949	0.956	0.959
	Power	0.571	0.856	0.956	0.992	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000
$\tilde{\beta}_3$	True Param. (e_p)	−0.37						1.09					
	Mean (m_p)	−0.377	−0.369	−0.369	−0.371	−0.369	−0.369	1.158	1.126	1.118	1.110	1.094	1.093
	Bias (peb)	0.018	−0.004	−0.002	0.004	−0.003	−0.003	0.062	0.033	0.026	0.018	0.004	0.003
	S.E. Mean (m_{se})	0.228	0.159	0.130	0.112	0.050	0.035	0.247	0.171	0.138	0.119	0.053	0.037
	S.E. Bias (seb)	0.013	0.026	0.013	0.026	0.033	0.038	0.005	0.014	0.001	0.048	0.003	0.009
	Coverage	0.951	0.946	0.948	0.941	0.936	0.941	0.957	0.949	0.956	0.954	0.948	0.948
	Power	0.383	0.633	0.806	0.904	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000
$\tilde{\beta}_4$	True Param. (e_p)	0.84						1.33					
	Mean (m_p)	0.866	0.859	0.853	0.847	0.842	0.842	1.399	1.361	1.348	1.340	1.335	1.332
	Bias (peb)	0.032	0.022	0.015	0.009	0.003	0.003	0.052	0.023	0.013	0.007	0.004	0.002
	S.E. Mean (m_{se})	0.291	0.203	0.165	0.143	0.063	0.045	0.313	0.217	0.176	0.152	0.068	0.048
	S.E. Bias (seb)	0.001	0.053	0.010	0.014	0.009	0.032	0.004	0.003	0.017	0.030	0.034	0.037
	Coverage	0.951	0.939	0.956	0.950	0.937	0.958	0.951	0.960	0.957	0.966	0.939	0.946
	Power	0.864	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 5

Results of the Monte Carlo simulation for covariances of the latent variables in the PIEG-2 model.

N_{obs}	Parameter	$Cov(\tilde{\eta}_1,\tilde{\eta}_2)$	$Cov(\tilde{\eta}_1,\tilde{\eta}_3)$	$Cov(\tilde{\eta}_2,\tilde{\eta}_3)$	$Cov(\tilde{\beta}_2,\tilde{\beta}_3)$	$Cov(\tilde{\beta}_2,\tilde{\beta}_4)$	$Cov(\tilde{\beta}_3,\tilde{\beta}_4)$	$Cov(\tilde{\beta}_2,\tilde{\eta}_1)$	$Cov(\tilde{\beta}_2,\tilde{\eta}_2)$	$Cov(\tilde{\beta}_2,\tilde{\eta}_3)$	$Cov(\tilde{\beta}_3,\tilde{\eta}_1)$	$Cov(\tilde{\beta}_3,\tilde{\eta}_2)$	$Cov(\tilde{\beta}_3,\tilde{\eta}_3)$	$Cov(\tilde{\beta}_4,\tilde{\eta}_1)$	$Cov(\tilde{\beta}_4,\tilde{\eta}_2)$	$Cov(\tilde{\beta}_4,\tilde{\eta}_3)$
	True Param. (e_p)	**1.00**	**0.80**	**1.17**	**0.15**	**1.15**	**0.29**	**0.10**	**0.13**	**0.03**	**-0.17**	**-0.11**	**-0.15**	**0.02**	**0.05**	**0.02**
250	Mean (m_p)	1.041	0.834	1.223	0.165	1.203	0.310	0.101	0.128	0.022	−0.187	−0.122	−0.160	0.016	0.046	0.022
	Bias (peb)	0.041	0.043	0.045	0.102	0.046	0.069	0.010	−0.015	−0.274	0.103	0.107	0.065	−0.220	−0.083	0.087
	S.E. Mean (m_{se})	0.314	0.302	0.335	0.158	0.229	0.173	0.225	0.240	0.236	0.216	0.231	0.227	0.246	0.261	0.257
	S.E. Bias (seb)	0.064	0.052	0.025	0.025	0.007	0.007	0.012	0.023	0.018	0.006	0.018	0.029	0.037	0.029	0.020
	Coverage	0.938	0.941	0.947	0.962	0.961	0.954	0.953	0.936	0.958	0.953	0.945	0.946	0.954	0.952	0.958
	Power	0.939	0.811	0.982	0.156	1.000	0.416	0.080	0.083	0.050	0.123	0.080	0.100	0.047	0.064	0.048
500	Mean (m_p)	1.020	0.816	1.196	0.159	1.173	0.302	0.094	0.125	0.022	−0.180	−0.118	−0.161	0.011	0.045	0.016
	Bias (peb)	0.020	0.020	0.022	0.060	0.020	0.042	−0.055	−0.042	−0.262	0.060	0.074	0.075	−0.436	−0.102	−0.182
	S.E. Mean (m_{se})	0.221	0.213	0.235	0.110	0.159	0.121	0.157	0.168	0.165	0.151	0.162	0.159	0.172	0.183	0.180
	S.E. Bias (seb)	0.067	0.030	0.029	0.007	0.014	0.005	0.024	0.024	0.021	0.005	0.001	0.025	0.025	0.017	0.030
	Coverage	0.933	0.946	0.944	0.950	0.953	0.952	0.944	0.945	0.938	0.955	0.954	0.944	0.937	0.938	0.947
	Power	0.997	0.984	1.000	0.296	1.000	0.708	0.099	0.126	0.063	0.204	0.104	0.159	0.060	0.058	0.058
750	Mean (m_p)	1.016	0.813	1.189	0.157	1.164	0.299	0.099	0.128	0.026	−0.178	−0.117	−0.157	0.020	0.049	0.016
	Bias (peb)	0.016	0.016	0.016	0.044	0.013	0.032	−0.011	−0.013	−0.148	0.049	0.067	0.045	−0.016	−0.019	−0.184
	S.E. Mean (m_{se})	0.180	0.173	0.192	0.089	0.129	0.099	0.128	0.137	0.134	0.123	0.132	0.129	0.140	0.149	0.146
	S.E. Bias (seb)	0.047	0.024	0.013	0.021	0.010	0.005	0.015	0.018	0.011	0.007	0.008	0.006	0.001	0.006	0.012
	Coverage	0.936	0.940	0.956	0.960	0.946	0.951	0.953	0.954	0.954	0.947	0.954	0.953	0.957	0.948	0.958
	Power	1.000	0.996	1.000	0.414	1.000	0.881	0.140	0.171	0.052	0.298	0.119	0.206	0.049	0.063	0.053
1000	Mean (m_p)	1.012	0.807	1.180	0.153	1.159	0.295	0.102	0.128	0.029	−0.177	−0.116	−0.154	0.020	0.049	0.019
	Bias (peb)	0.012	0.009	0.009	0.020	0.008	0.017	0.019	−0.012	−0.034	0.039	0.054	0.027	−0.002	−0.015	−0.068
	S.E. Mean (m_{se})	0.155	0.149	0.165	0.077	0.112	0.085	0.110	0.118	0.116	0.106	0.114	0.112	0.121	0.128	0.126
	S.E. Bias (seb)	0.062	0.040	0.000	0.036	0.043	0.029	0.001	0.021	0.013	0.027	0.022	0.006	0.022	0.004	0.014
	Coverage	0.930	0.942	0.952	0.959	0.950	0.953	0.953	0.953	0.945	0.961	0.956	0.943	0.954	0.962	0.953
	Power	1.000	1.000	1.000	0.500	1.000	0.959	0.161	0.207	0.063	0.381	0.164	0.262	0.053	0.066	0.046
5000	Mean (m_p)	1.003	0.798	1.170	0.151	1.153	0.291	0.103	0.131	0.029	−0.171	−0.113	−0.150	0.022	0.049	0.020
	Bias (peb)	0.003	−0.002	0.000	0.005	0.003	0.002	0.025	0.004	−0.017	0.004	0.026	0.003	0.085	−0.023	0.007
	S.E. Mean (m_{se})	0.069	0.067	0.074	0.034	0.050	0.038	0.049	0.053	0.052	0.047	0.051	0.050	0.054	0.057	0.056
	S.E. Bias (seb)	0.011	0.021	0.006	0.013	0.004	0.011	0.018	0.004	0.014	0.020	0.032	0.010	0.021	0.019	0.026
	Coverage	0.952	0.949	0.948	0.947	0.950	0.950	0.948	0.948	0.948	0.946	0.956	0.949	0.950	0.946	0.934
	Power	1.000	1.000	1.000	0.995	1.000	1.000	0.541	0.692	0.095	0.952	0.586	0.865	0.078	0.130	0.073
10000	Mean (m_p)	1.003	0.801	1.169	0.151	1.152	0.290	0.102	0.130	0.029	−0.170	−0.112	−0.150	0.022	0.050	0.021
	Bias (peb)	0.003	0.001	−0.001	0.005	0.002	0.000	0.021	0.002	−0.027	0.000	0.017	0.002	0.096	0.000	0.036
	S.E. Mean (m_{se})	0.049	0.047	0.052	0.024	0.035	0.027	0.035	0.037	0.037	0.033	0.036	0.035	0.038	0.041	0.040
	S.E. Bias (seb)	0.030	0.006	0.021	0.018	0.001	0.002	0.023	0.031	0.012	0.008	0.024	0.004	0.028	0.013	0.012
	Coverage	0.950	0.950	0.946	0.939	0.949	0.948	0.954	0.943	0.945	0.951	0.937	0.955	0.949	0.942	0.946
	Power	1.000	1.000	1.000	1.000	1.000	1.000	0.846	0.939	0.135	1.000	0.877	0.994	0.100	0.224	0.090

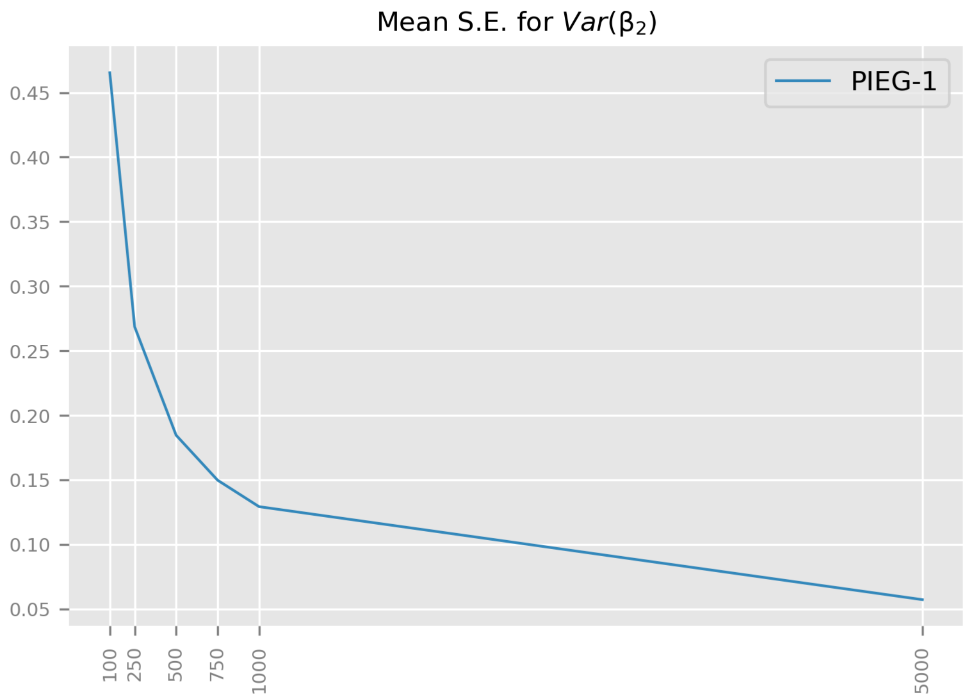

Figure 2. Mean standard errors for $Var(\tilde{\beta}_2)$ across all samples generated in each of the the Monte Carlo simulations for PIEG-1.

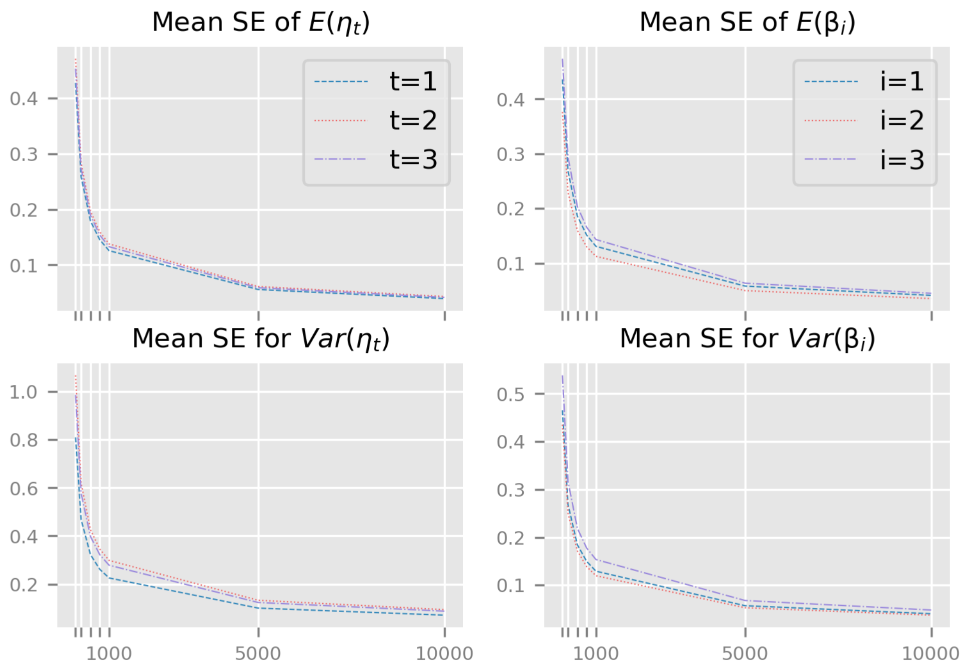

Figure 3. Mean standard errors across all samples generated in each of the Monte Carlo simulations for PIEG-1.

A.2. M*plus* inputs

Listing 1: Monte Carlo simulation M*plus* input.

```
TITLE:  Montecarlo PIEG Model 1

MONTECARLO:
NAMES ARE y1-y12;
NOBSERVATIONS = 250;
NREPS = 1000;
GENERATE = y1-y12 (4 p);
POPULATION = m1_estimates_r.dat;
COVERAGE = m1_estimates_r.dat;
CATEGORICAL ARE y1-y12;

ANALYSIS:
ESTIMATOR = WLSMV;
PARAMETERIZATION = Theta;


MODEL POPULATION:

ETA1 BY y1-y4@1;
ETA2 BY y5-y8@1;
ETA3 BY y9-y12@1;

BETA2  BY y2@1 y6@1 y10@1;
BETA3  BY y3@1 y7@1 y11@1;
BETA4  BY y4@1 y8@1 y12@1;

[ETA1*];[ETA2*];[ETA3*];
[BETA2*];[BETA3*];[BETA4*];

[y1$1@0];            [y5$1@0];            [y9$1@0];
[y1$2*] (t1i1k2_p); [y5$2*] (t2i1k2_p); [y9$2*] (t3i1k2_p);
```

```
[y1$3*] (t1i1k3_p); [y5$3*] (t2i1k3_p); [y9$3*] (t3i1k3_p);
[y1$4*] (t1i1k4_p); [y5$4*] (t2i1k4_p); [y9$4*] (t3i1k4_p);

[y2$1@0];           [y6$1@0];           [y10$1@0];
[y2$2*] (t1i2k2_p); [y6$2*] (t2i2k2_p); [y10$2*] (t3i2k2_p);
[y2$3*] (t1i2k3_p); [y6$3*] (t2i2k3_p); [y10$3*] (t3i2k3_p);
[y2$4*] (t1i2k4_p); [y6$4*] (t2i2k4_p); [y10$4*] (t3i2k4_p);

[y3$1@0];           [y7$1@0];           [y11$1@0];
[y3$2*] (t1i3k2_p); [y7$2*] (t2i3k2_p); [y11$2*] (t3i3k2_p);
[y3$3*] (t1i3k3_p); [y7$3*] (t2i3k3_p); [y11$3*] (t3i3k3_p);
[y3$4*] (t1i3k4_p); [y7$4*] (t2i3k4_p); [y11$4*] (t3i3k4_p);

[y4$1@0];           [y8$1@0];           [y12$1@0];
[y4$2*] (t1i4k2_p); [y8$2*] (t2i4k2_p); [y12$2*] (t3i4k2_p);
[y4$3*] (t1i4k3_p); [y8$3*] (t2i4k3_p); [y12$3*] (t3i4k3_p);
[y4$4*] (t1i4k4_p); [y8$4*] (t2i4k4_p); [y12$4*] (t3i4k4_p);


MODEL:

ETA1 BY y1-y4@1;
ETA2 BY y5-y8@1;
ETA3 BY y9-y12@1;

BETA2  BY y2@1 y6@1 y10@1;
BETA3  BY y3@1 y7@1 y11@1;
BETA4  BY y4@1 y8@1 y12@1;

[ETA1*];[ETA2*];[ETA3*];
[BETA2*];[BETA3*];[BETA4*];

[y1$1@0];         [y5$1@0];         [y9$1@0];
[y1$2*] (t1i1k2); [y5$2*] (t2i1k2); [y9$2*] (t3i1k2);
[y1$3*] (t1i1k3); [y5$3*] (t2i1k3); [y9$3*] (t3i1k3);
[y1$4*] (t1i1k4); [y5$4*] (t2i1k4); [y9$4*] (t3i1k4);

[y2$1@0];         [y6$1@0];         [y10$1@0];
[y2$2*] (t1i2k2); [y6$2*] (t2i2k2); [y10$2*] (t3i2k2);
[y2$3*] (t1i2k3); [y6$3*] (t2i2k3); [y10$3*] (t3i2k3);
[y2$4*] (t1i2k4); [y6$4*] (t2i2k4); [y10$4*] (t3i2k4);

[y3$1@0];         [y7$1@0];         [y11$1@0];
[y3$2*] (t1i3k2); [y7$2*] (t2i3k2); [y11$2*] (t3i3k2);
[y3$3*] (t1i3k3); [y7$3*] (t2i3k3); [y11$3*] (t3i3k3);
[y3$4*] (t1i3k4); [y7$4*] (t2i3k4); [y11$4*] (t3i3k4);

[y4$1@0];         [y8$1@0];         [y12$1@0];
[y4$2*] (t1i4k2); [y8$2*] (t2i4k2); [y12$2*] (t3i4k2);
[y4$3*] (t1i4k3); [y8$3*] (t2i4k3); [y12$3*] (t3i4k3);
[y4$4*] (t1i4k4); [y8$4*] (t2i4k4); [y12$4*] (t3i4k4);
```

Listing 2: PIEG-1 model M*plus* input.

```
TITLE:      PIEG Model 1;

VARIABLE:
NAMES ARE y1-y12;
USEVARIABLES ARE  y1-y12 ;
CATEGORICAL ARE  y1-y12 ;

ANALYSIS:
Estimator = WLSMV;
PARAMETERIZATION=THETA;


MODEL:

ETA1 BY y1-y4@1;
ETA2 BY y5-y8@1;
ETA3 BY y9-y12@1;

BETA2  BY y2@1 y6@1 y10@1;
BETA3  BY y3@1 y7@1 y11@1;
BETA4  BY y4@1 y8@1 y12@1;

[ETA1*];[ETA2*];[ETA3*];
[BETA2*];[BETA3*];[BETA4*];

[y1$1@0];         [y5$1@0];         [y9$1@0];
[y1$2*] (t1i1k2); [y5$2*] (t2i1k2); [y9$2*] (t3i1k2);
[y1$3*] (t1i1k3); [y5$3*] (t2i1k3); [y9$3*] (t3i1k3);
[y1$4*] (t1i1k4); [y5$4*] (t2i1k4); [y9$4*] (t3i1k4);
```

```
[y2$1@0];          [y6$1@0];          [y10$1@0];
[y2$2*] (t1i2k2); [y6$2*] (t2i2k2); [y10$2*] (t3i2k2);
[y2$3*] (t1i2k3); [y6$3*] (t2i2k3); [y10$3*] (t3i2k3);
[y2$4*] (t1i2k4); [y6$4*] (t2i2k4); [y10$4*] (t3i2k4);

[y3$1@0];          [y7$1@0];          [y11$1@0];
[y3$2*] (t1i3k2); [y7$2*] (t2i3k2); [y11$2*] (t3i3k2);
[y3$3*] (t1i3k3); [y7$3*] (t2i3k3); [y11$3*] (t3i3k3);
[y3$4*] (t1i3k4); [y7$4*] (t2i3k4); [y11$4*] (t3i3k4);

[y4$1@0];          [y8$1@0];          [y12$1@0];
[y4$2*] (t1i4k2); [y8$2*] (t2i4k2); [y12$2*] (t3i4k2);
[y4$3*] (t1i4k3); [y8$3*] (t2i4k3); [y12$3*] (t3i4k3);
[y4$4*] (t1i4k4); [y8$4*] (t2i4k4); [y12$4*] (t3i4k4);
```

Listing 3: PIEG-2 model M*plus* input.

```
TITLE:      PIEG Model 2;

VARIABLE:
NAMES ARE y1-y12;
USEVARIABLES ARE  y1-y12 ;
CATEGORICAL ARE  y1-y12 ;

ANALYSIS:
Estimator = WLSMV;
PARAMETERIZATION=THETA;


MODEL:

ETA1 BY y1-y4@1;
ETA2 BY y5-y8@1;
ETA3 BY y9-y12@1;

BETA2  BY y2@1 y6@1 y10@1;
BETA3  BY y3@1 y7@1 y11@1;
BETA4  BY y4@1 y8@1 y12@1;

[ETA1*];[ETA2*];[ETA3*];
[BETA2*];[BETA3*];[BETA4*];

[y1$1@0];          [y5$1@0];          [y9$1@0];
[y1$2*] (t1i1k2); [y5$2*] (t1i1k2); [y9$2*] (t1i1k2);
[y1$3*] (t1i1k3); [y5$3*] (t1i1k3); [y9$3*] (t1i1k3);
[y1$4*] (t1i1k4); [y5$4*] (t1i1k4); [y9$4*] (t1i1k4);

[y2$1@0];          [y6$1@0];          [y10$1@0];
[y2$2*] (t1i2k2); [y6$2*] (t1i2k2); [y10$2*] (t1i2k2);
[y2$3*] (t1i2k3); [y6$3*] (t1i2k3); [y10$3*] (t1i2k3);
[y2$4*] (t1i2k4); [y6$4*] (t1i2k4); [y10$4*] (t1i2k4);

[y3$1@0];          [y7$1@0];          [y11$1@0];
[y3$2*] (t1i3k2); [y7$2*] (t1i3k2); [y11$2*] (t1i3k2);
[y3$3*] (t1i3k3); [y7$3*] (t1i3k3); [y11$3*] (t1i3k3);
[y3$4*] (t1i3k4); [y7$4*] (t1i3k4); [y11$4*] (t1i3k4);

[y4$1@0];          [y8$1@0];          [y12$1@0];
[y4$2*] (t1i4k2); [y8$2*] (t1i4k2); [y12$2*] (t1i4k2);
[y4$3*] (t1i4k3); [y8$3*] (t1i4k3); [y12$3*] (t1i4k3);
[y4$4*] (t1i4k4); [y8$4*] (t1i4k4); [y12$4*] (t1i4k4);
```