

*Center for Advanced Studies in  
Measurement and Assessment*

*CASMA Research Report*

*Number 15*

**Bootstrap Techniques for  
Estimating Variability in  
Generalizability Theory**

*Ye Tong and Robert L. Brennan*

March 2006

Center for Advanced Studies in  
Measurement and Assessment (CASMA)  
College of Education  
University of Iowa  
Iowa City, IA 52242  
Tel: 319-335-5439  
Fax: 319-384-0505  
Web: [www.education.uiowa.edu/casma](http://www.education.uiowa.edu/casma)

All rights reserved

## Contents

<b>Abstract</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Methodology</b>	<b>1</b>
2.1 Data Simulation . . . . .	1
2.1.1 Normal Data . . . . .	1
2.1.2 Dichotomous Data . . . . .	2
2.1.3 Polytomous Data . . . . .	3
2.2 Bootstrap . . . . .	3
2.3 Bias Correction . . . . .	4
<b>3 Results</b>	<b>5</b>
3.1 $p \times i \times h$ Design . . . . .	5
3.1.1 Estimated Variance Components . . . . .	6
3.1.2 Estimated Standard Errors . . . . .	6
3.2 $p \times (i:h)$ Design . . . . .	7
3.2.1 Estimated Variance Components . . . . .	7
3.2.2 Estimated Standard Errors . . . . .	7
<b>4 Discussion</b>	<b>7</b>
<b>5 References</b>	<b>9</b>

## List of Tables

1	Parameter Values for Variance Components and Standard Errors for the Three Types of Data. . . . .	10
2	Estimated Variance Components for the $p \times i \times h$ Design, Raw Bootstrap Results. . . . .	11
3	Estimated Variance Components for the $p \times i \times h$ Design, Bias Adjusted Bootstrap Results. . . . .	12
4	Estimated Standard Errors for the $p \times i \times h$ Design, Raw Bootstrap Results . . . . .	13
5	Estimated Standard Errors for the $p \times i \times h$ Design, Bias Adjusted Bootstrap Results. . . . .	14
6	Estimated Variance Components for the $p \times (i:h)$ Design, Raw Bootstrap Results. . . . .	15
7	Estimated Variance Components for the $p \times (i:h)$ Design, Biased Adjusted Bootstrap Results. . . . .	16
8	Estimated Standard Errors for the $p \times (i:h)$ Design, Raw Bootstrap Results. . . . .	17
9	Estimated Standard Errors for the $p \times (i:h)$ Design, Biased Adjusted Bootstrap Results. . . . .	18

## Abstract

The estimation of standard errors of estimated variance components has long been a challenging task in generalizability theory. For two decades researchers have speculated about the potential applicability of the bootstrap for obtaining such estimates, but the same researchers have noted that bootstrap procedures give biased estimates of variance components and their standard errors. This paper proposes general bias-correction formulas that may ameliorate these problems. Although no rigorous derivations are provided, simulation results suggest that the bias correction formulas work reasonably well.

## 1 Introduction

One major element in generalizability theory is the estimation of variance components. Like other statistics, variance components are subject to sampling error. The estimation of standard errors of estimated variance components has long been a challenging task. Some possible approaches include a normality-based derivation, nonparametric-based bootstrap and jackknife procedures, and Bayesian-based MCMC approaches.

The bootstrap procedure has shown some promises due to its straightforward applicability. This technique, however, is also beset with potential complexity. Some of these problems are associated with the fact that bootstrap procedures generally produce biased estimates of variance components and their standard errors (Brennan, 2001; Wiley, 2000). Other problems are associated with the fact that, with ANOVA-type designs, there are many possibilities for obtaining bootstrap samples, and it has not been demonstrated that there is a single optimal bootstrap procedure for obtaining the “best” estimates for all variance components and standard errors.

Previous research has shown that different bootstrap sampling strategies provide widely divergent point and interval estimates for variance components and their standard errors (Brennan et al., 1987; Luecht & Smith, 1988; Othman, 1995; Wiley, 2000; Brennan, 2001; Tong & Brennan, 2004). Furthermore, these papers allude to sample size and disproportionality between the numbers of persons and items as possible factors contributing to the differences across bootstrap strategies. Wiley (2000) provided algorithms explaining why the discrepancies exist and how to correct them in the  $p \times i$  context. In this paper, bias correction factors for applying bootstrap techniques are proposed for more complex designs and are empirically examined in the context of simulated normal, dichotomous and polytomous data for two multi-facet designs:  $p \times i \times h$  and  $p \times (i:h)$ . An example of the  $p \times i \times h$  design occurs when examinees ( $p$ ) take all test items ( $i$ ) and all items are scored by the same raters ( $h$ ). An example of the  $p \times (i:h)$  design occurs when examinees take all items, but raters only score a subset of items.

## 2 Methodology

In this section, the data simulation procedure is presented first, followed by the proposed bias correction factors for the bootstrap application in estimating standard errors of variance components. The sample sizes used were:  $n_p = 100$ ,  $n_i = 20$  and  $n_h = 2$  for both designs. That is, 100 person, 20 items and 2 raters.

### 2.1 Data Simulation

#### 2.1.1 Normal Data

For the normality-based procedure, score effects are assumed to be normally distributed. For the  $p \times i \times h$  design, the following formula was used to generate

observed scores:

$$X_{pjh} = \mu + \sigma(p)z_p + \sigma(i)z_i + \sigma(h)z_h + \sigma(pi)z_{pi} + \sigma(ph)z_{ph} + \sigma(ih)z_{ih} + \sigma(pih)z_{pih}, \quad (1)$$

where  $z_p, z_i, z_h, z_{pi}, z_{ph}, z_{ih}$  and  $z_{pih}$  were randomly and independently sampled from the standard normal distribution. Similarly, to generate observed scores for the  $p \times (i:h)$  design, the following formula was used:

$$X_{pi:h} = \mu + \sigma(p)z_p + \sigma(h)z_h + \sigma(ph)z_{ph} + \sigma(i:h)z_{i:h} + \sigma(pi:h)z_{pi:h}, \quad (2)$$

with the  $z$  scores similarly defined.

Based on the random-effects assumption, the actual standard errors of the estimated variance components,  $SE(\alpha)$  (where  $\alpha$  is used generically for a variance component) can be obtained using the following formula (see Searle, 1971) referenced in Brennan (2001):

$$\sigma[\hat{\sigma}^2(\alpha)] = \frac{1}{\pi(\hat{\alpha})} \sqrt{\sum_{\beta} \frac{2[\mathbf{EMS}(\beta)]^2}{df(\beta)}}, \quad (3)$$

where  $\pi(\hat{\alpha})$  is the product of the sample sizes for indices not in  $\alpha$ , and  $\beta$  indexes the expected mean squares that enter  $\sigma^2(\alpha)$ . For example, under normality, the standard error of  $\hat{\sigma}^2(i:h)$  for the nested  $p \times (i:h)$  design is:

$$\sigma(\hat{\sigma}^2(i:h)) = \frac{1}{n_p} \sqrt{\frac{2[\mathbf{EMS}(i:h)]^2}{n_h(n_i - 1)} + \frac{2[\mathbf{EMS}(pi:h)]^2}{n_h(n_p - 1)(n_i - 1)}}.$$

See Table 1 for the parameter values for variance components and standard errors used in this study for the normal data. These values were chosen to represent typical relative magnitudes of the variance components in these two designs.

### 2.1.2 Dichotomous Data

To simulate dichotomous data, Equations 1 and 2 were again used for the two designs, respectively. If the simulated score exceeds 1, a response of 1 was assigned; otherwise, a response of 0 was assigned to obtain dichotomous type of data.

Based on the above simulation approach, the parameter values for variance components and standard errors for dichotomous data were not readily available. To obtain reasonable substitutes, 5000 such datasets were simulated, with each dataset consisting of  $n_p = 100$ ,  $n_i = 20$  and  $n_h = 2$ . Variance components were estimated based on each of the 5000 datasets. The means of these variance components were treated as the parameter values for variance components; the

standard deviations of these estimated variance components were considered the standard errors for the variance components. Table 1 reports these “parameter values” for dichotomous data. As can be observed from Table 1, the relative magnitudes of the variance components and standard errors for dichotomous data have similar trends to that observed for normal data.

### 2.1.3 Polytomous Data

If  $B$  stands for a randomly and independently selected binomial value, Feng (2002, p.77) showed that the following formulas can be used to simulate polytomous data for the  $p \times i \times h$  and  $p \times (i:h)$  designs, respectively:

$$X_{pih} = B_p + B_i + B_h + B_{pi} + B_{ph} + B_{ih} + B_{pih}, \quad (4)$$

and

$$X_{pi:h} = B_p + B_h + B_{i:h} + B_{ph} + B_{pi:h}. \quad (5)$$

The sample sizes and parameter values for the simulation are reported in Table 1. The parameters for the  $p \times i \times h$  design were specified based on results obtained by Gao and Brennan (2001); parameters for the  $p \times (i:h)$  design were based on Lane, Liu, Ankenmann, and Stone (1996). Those two studies both provided variance component estimates for multiple forms, and the averages of those multiple forms are the values reported in Table 1. The scores range from 0 to 10 for each item in both designs. The following equations were used to simulate the data for the two designs:

$$\begin{aligned} X = & \text{BIN}(2, .7966) + \text{BIN}(1, .8570) + \text{BIN}(1, .9879) + \text{BIN}(2, .7313) \\ & + \text{BIN}(1, .9858) + \text{BIN}(1, .9975) + \text{BIN}(2, .8025), \end{aligned}$$

and

$$X = \text{BIN}(1, .713) + \text{BIN}(1, .843) + \text{BIN}(2, .713) + \text{BIN}(1, .930) + \text{BIN}(5, .630),$$

where  $\text{BIN}(2, .7966)$  means a binomial value with number of trials being 2 and the probability of a success for each trial being .7966. Other binomial values are defined in the same way.

Similar to the dichotomous data case, no formula exists for calculating the parameters for standard errors for polytomous data. Again, to obtain reasonable substitutes, 5000 datasets were generated for each design using the above equations. For each of the 5000 datasets, variance components were estimated and the standard deviations of these estimated variance components were treated as the parameter values for standard errors (See Table 1).

## 2.2 Bootstrap

As previously mentioned, for any given design, a number of different bootstrap samples can be extracted. For example, with respect to the  $p \times i \times h$  design,

examinees can be the only facet<sup>1</sup> resampled (boot- $p$ ), items can be the only facet resampled (boot- $i$ ), raters can be the only facet resampled (boot- $h$ ), or any of the combinations of the three facets can be resampled. In fact, any of the following seven bootstrap procedures can be employed: boot- $p$ , boot- $i$ , boot- $h$ , boot- $p, i$ , boot- $p, h$ , boot- $i, h$  and boot- $p, i, h$ . For example, boot- $p, i$  involves resampling of persons and items and matching the sampled persons with the sampled items.

Similarly, for the  $p \times (i:h)$  design, any of the following five bootstrap procedures can be employed: boot- $p$ , boot- $h$ , boot- $i, h$ , boot- $p, h$ , and boot- $p, i, h$ . For the  $p \times (i:h)$  design, because the  $i$  facet is nested with the  $h$  facet, the bootstrap technique is a bit more complicated. For example, to conduct the boot- $i, h$  procedure, an independent sample of size  $n_h$ , with replacement, is drawn first. Within each drawn  $h$ , a sample of size  $n_i$  is drawn with replacement.

### 2.3 Bias Correction

Wiley (2000) demonstrated that bootstrap estimates are generally biased in estimating variance components and standard errors. He formulated bias correction factors for  $p \times i$  design. Here bias correction factors for more complex designs are presented. Let  $\sigma^2(\alpha)$  be the variance component of interest, and let  $K$  be the set of facets in the bootstrap sampling plan, with each of them designated  $k_j$ . If  $K$  and  $\alpha$  contain common facet(s), designate them  $c_i$ . Then the following bias correction factors are proposed:

$$\begin{aligned} \hat{\sigma}^2(\alpha) &= \frac{\prod_{c_i} n_{c_i}}{\prod_{c_i} df(c_i)} \hat{\sigma}^2(\alpha | \text{boot-}K) \\ &\quad - \sum_{k_j \notin \alpha} \frac{\prod_{c_i} n_{c_i}}{df(k_j) \prod_{c_i} df(c_i)} \hat{\sigma}^2(\alpha k_j | \text{boot-}K) \\ &\quad + \sum_{\substack{k_j \notin \alpha, k_m \notin \alpha \\ k_j \neq k_m}} \frac{n_{c_i}}{df(k_m) df(k_j) df(c_i)} \hat{\sigma}^2(\alpha k_j k_m | \text{boot-}K). \end{aligned} \quad (6)$$

If  $K$  and  $\alpha$  do not contain any common facet, then the following bias correction factors are used:

$$\begin{aligned} \hat{\sigma}^2(\alpha) &= \hat{\sigma}^2(\alpha | \text{boot-}K) \\ &\quad - \sum_{k_j} \frac{1}{df(k_j)} \hat{\sigma}^2(\alpha k_j | \text{boot-}K) \end{aligned}$$

<sup>1</sup>Technically, in generalizability theory,  $p$  is usually called the objects of measurement rather than a facet.

$$+ \sum_{k_j \neq k_m} \frac{1}{df(k_m)df(k_j)} \hat{\sigma}^2(\alpha k_j k_m | \text{boot} - K). \quad (7)$$

Note that in Equation 6, if there are no  $c_i$  indices, then terms involving them are set to 1. For example, to estimate  $\hat{\sigma}^2(p)$  under the  $p \times i \times h$  design using the boot- $i, h$  procedure,  $\alpha$  is  $p$  and  $K$  consists of  $i$  and  $h$ . In this case,  $\alpha$  and  $K$  do not contain a common facet, and Equation 7 can be used to obtain the following correction:

$$\begin{aligned} \hat{\sigma}^2(p) &= \hat{\sigma}^2(p | \text{boot} - i, h) - \frac{1}{n_i - 1} \hat{\sigma}^2(pi | \text{boot} - i, h) \\ &\quad - \frac{1}{n_h - 1} \hat{\sigma}^2(ph | \text{boot} - i, h) + \frac{1}{(n_i - 1)(n_h - 1)} \hat{\sigma}^2(pih | \text{boot} - i, h). \end{aligned}$$

A complete list of bias correction factors by bootstrap strategy can be found in Tong and Brennan (2004).

In most applications in generalizability theory, estimation of relative errors [ $\sigma^2(\delta)$ ] and absolute errors [ $\sigma^2(\Delta)$ ] are of major concern. The standard errors of estimated relative and absolute errors were also calculated in this study for different bootstrap samples. The parameter values for the standard errors of  $\sigma^2(\delta)$  and  $\sigma^2(\Delta)$  [ $SE(\delta)$  and  $SE(\Delta)$ ] were substituted with the standard deviations over the 5000 trials for dichotomous and polytomous data, as reported in Table 1.

### 3 Results

The simulation conducted involved 1000 trials. That is, 1000 arrays of size  $n_p \times n_i \times n_h$  ( $100 \times 20 \times 2$ ) were generated. For each bootstrap procedure [7 for the  $p \times i \times h$  design and 5 for the  $p \times (i:h)$  design], 100 bootstrap samples (replications) were drawn within each of the 1000 trials. That is, the results for each of the bootstrap procedures for both designs were based on 100,000 bootstrap analysis. Within each trial, for each effect, the mean and variance, over the 100 replications, of the estimated variance components (both adjusted using the bias corrections and unadjusted) were obtained. For each effect, the average over trials of the mean estimated variance component (both unadjusted and adjusted) was taken as the estimated variance component for the simulation; similarly, the square root of the variances of the estimated variance components averaged over trials (both unadjusted and adjusted) was taken as the estimated standard error for the simulation. A C program was written to conduct all the simulation and the analysis.

#### 3.1 $p \times i \times h$ Design

As described in the methodology section, normal, dichotomous and polytomous data were generated. All seven bootstrap procedures for this design were carried

out and both variance components and standard errors were estimated using raw bootstrap results and adjusted bootstrap results. Tables 2 and 3 report estimated variance components from the bootstrap procedures and Tables 4 and 5 report estimated standard errors.

### 3.1.1 Estimated Variance Components

Table 2 reports the original bootstrap variance component estimates and Table 3 reports the bias-adjusted results. For each of the tables, results for all three types of data are reported. For each type of data, parameter values for the variance components are presented first, followed by the results for each of the seven bootstrap procedures. As mentioned in the previous section, each of the estimates in the tables was based on 100,000 bootstrap calculations (100 bootstrap estimates within each of the 1000 trials). As can be observed from Table 2, however, some of the estimates deviated substantially from their corresponding parameter values. For example, with normal data, the parameter value for  $\sigma^2(pih)$  was 144, but the estimate using the *boot- $i, h$*  procedure yielded an estimate of 68.4561, even though the estimate was based on 100,000 bootstrap calculations. The results in Table 2 empirically demonstrated that the raw results from bootstrap procedures are generally biased.

After applying the bootstrap adjustment proposed in this paper, however, the estimates across all bootstrap procedures for all three types of data became very similar to their corresponding parameter values, as can be observed from Table 3. Following the previous example, the *boot- $i, h$*  estimate of  $\sigma^2(pih)$  became 144.1182, much closer to the parameter value of 144 than before. It appears that the bias correction factors worked reasonably well for estimating variance components.

### 3.1.2 Estimated Standard Errors

Tables 4 and 5 provide estimated results of standard errors for estimated variance components, with Table 4 reporting raw bootstrap results and Table 5 presenting bias-adjusted bootstrap results. These two tables have the same layout as Tables 2 and 3, reporting parameter values in the first row for each data type, followed by estimated standard errors for all seven bootstrap procedures. Again, the raw bootstrap values in Table 4 appeared to be quite different from their corresponding parameters. For example, the standard error estimate of  $\sigma^2(pih)$  for *boot- $i, h$*  was 68.5026, whereas its parameter value was only 4.6955.

Table 5 reports results based on the bias adjustments. These results are mixed. Some of the bootstrap estimates were made close their parameters, such as *boot- $p$*  estimates of  $SE(p)$ ,  $SE(pi)$ ,  $SE(ph)$  and  $SE(pih)$ . However, there are some estimates of standard errors that deviated quite substantially from the parameters, such as *boot- $i, h$*  estimates for just about all  $SE$ 's. It appears that across all three data types, *boot- $p$*  produced the most adequate results for  $SE(p)$ ,  $SE(pi)$ ,  $SE(ph)$  and  $SE(pih)$ ; *boot- $i$*  produced the most adequate results for  $SE(i)$  and  $SE(ih)$ ; *boot- $h$*  produced the most adequate results for

$SE(h)$ . None of the bootstrap procedures consistently produced good estimates for  $SE(\Delta)$ . Note, in particular, that boot- $p$  consistently underestimated the standard error for estimated absolute error variance, although its estimate of the magnitude of the absolute error variance was fairly accurate.

### 3.2 $p \times (i:h)$ Design

Tables 6 through 9 report estimates obtained with this nested design, across all three data types. The layout of the tables is similar to the crossed design, with Tables 6 and 7 reporting raw and adjusted estimates for variance components and Tables 8 and 9 reporting raw and adjusted estimates for standard errors.

#### 3.2.1 Estimated Variance Components

Similar to the crossed design, raw bootstrap estimates (Table 6) are quite different from the parameter values, even though each of the estimates was based on 100,000 bootstrap calculations. After bias adjustment (Table 7), however, unlike the crossed design, not all the estimates were made similar to their corresponding parameters, especially with the boot- $h$  and boot- $ph$  procedures, where, for example, there were some negative values for variance component estimates. By contrast, the results for boot- $p$ , boot- $i, h$  and boot- $p, i, h$  were reasonably close to the parameters. It appeared that with the nested design, the bias adjustment worked well with some bootstrap procedures, but not all of them.

#### 3.2.2 Estimated Standard Errors

Similar to what was observed with the crossed design, the bias adjusted results improved upon the raw results in terms of accuracy, but not as well as for the crossed design. Similar to the crossed design, boot- $p$  produced the most accurate results for  $SE(p)$ ,  $SE(ph)$ ,  $SE(pi:h)$  and  $SE(\delta)$ ; boot- $h$  yielded the most adequate results for  $SE(h)$  and boot- $i, h$  produced the most accurate results for  $SE(i:h)$ . Again, no bootstrap procedure appeared to produce an accurate estimate for  $SE(\Delta)$  across all three data type. Boot- $p$  consistently underestimated  $SE(\Delta)$ , but produced accurate estimates for  $\sigma^2(\Delta)$ .

## 4 Discussion

For multi-facet designs, there are many possibilities for obtaining bootstrap samples, and it has not been demonstrated that there is a single optimal bootstrap procedure for obtaining the “best” estimates for all variance components and standard errors. Studies have shown that bootstrap estimates in these contexts are generally biased. Bias correction factors are proposed in this paper. Using simulated normal, dichotomous and polytomous data, various bootstrap procedures were investigated for the  $p \times i \times h$  and  $p \times (i:h)$  designs. The results from the study seem to warrant the following tentative conclusions:

1. Bootstrap procedures for estimating variance components and standard errors generally yield bias results.
2. The bias correction formulas proposed in this paper (Equations 6 and 7) performed reasonably well with the three simulated data types, especially for the  $p \times i \times h$  design.
3. Specifically, for main effects, the corresponding bootstrap procedure generally yields good results [e.g., boot- $i$  to estimate  $\sigma^2(i)$  and  $SE(i)$ ]. For interaction effects, the one-dimensional bootstrap procedure corresponding to the facet in the interaction term with the largest sample size generally produces adequate results [e.g., boot- $p$  to estimate  $\sigma^2(ph)$  and  $SE(ph)$ ].
4. The boot- $p$  procedure consistently yielded the most accurate results for estimating  $SE(\delta)$ .
5. No bootstrap procedure performed well for estimating  $SE(\Delta)$ ; however, boot- $p$  performed quite well for estimating the magnitude of  $\sigma^2(\Delta)$ . Although not discussed in detail here, when  $n_h$  was 4, 6, and 8, boot- $p$  performed very well for estimating the value of  $\sigma^2(\Delta)$ , with estimated  $SE(\Delta)$  close to 0.

The estimation of standard errors of estimated variance components has long been a challenging task in generalizability theory. For two decades researchers have speculated about the potential applicability of the bootstrap for obtaining such estimates, but the same researchers have identified problems in using the bootstrap. This paper proposes procedures that may ameliorate these problems and thus make the bootstrap a more accessible procedure for estimating standard errors of estimated variance components. No rigorous derivations were provided for the bias correction formulas; however, the simulation results suggest that the bias correction formulas worked reasonably well.

## 5 References

- Brennan, R. L., Harris, D.J., & Hanson, B.A. (1987). *The bootstrap and other procedures for examining the variability of estimated variance components in testing contexts*. ACT Research Report Series 87-7. Iowa City, IA: American College Testing Program.
- Brennan, R. L. (2001). *Generalizability Theory*. New York:Springer-Verlag.
- Feng, W. C., & Ankenmann, R. D. (2003) *Applicability of the jackknife procedure for estimating standard errors of variance component estimates in random effects i:p and px(i:r) G study designs*. Paper presented at the Annual Meeting of National Council on Measurement in Education, Chicago, IL.
- Gao, X., & Brennan, R. L. (2001). Variability of estimated variance components and related statistics in a performance assessment. *Applied Measurement in Education*, *14*(2), 191–203.
- Lane, S., Liu, M., Ankenmann, R. D., & Stone, C. A. (1996). Generalizability and validity of a mathematics performance assessment. *Journal of Educational Measurement*, *33*(1), 71–92.
- Leucht, R.M., & Smith, P.L.(1989). *The effects of bootstrapping strategies on the estimation of variance components*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Othman, A.R. (1995). *Examining task sampling variability in science performance assessments*. Unpublished doctoral dissertation, University of California, Santa Barbara.
- Searle, S. R. (1971). *Linear Models*. New York: Wiley.
- Tong, Y., & Brennan, R. L. (2004). *Bootstrap strategies for estimating standard errors of estimated variance components*. CASMA Research Report Number 4, University of Iowa, Iowa City, IA.
- Wiley, E. W. (2000). *Bootstrap strategies for variance component estimation: theoretical and empirical results*. Unpublished doctoral dissertation, Stanford.

Table 1: Parameter Values for Variance Components and Standard Errors for the Three Types of Data.

Sample Size	$p \times i \times h$			$p \times (i:h)$		
	Variance Component	Standard Error <sup>a</sup>		Variance Component	Standard Error	
<b>Normal Data</b>						
$n_p = 100$	$\sigma^2(p)$	16.0000	3.4545	$\sigma^2(p)$	16.0000	3.2761
$n_i = 20$	$\sigma^2(i)$	4.0000	2.3399	$\sigma^2(h)$	1.0000	2.0872
$n_h = 2$	$\sigma^2(h)$	1.0000	1.7580	$\sigma^2(ih)$	7.0000	2.0836
	$\sigma^2(pi)$	64.0000	5.0178	$\sigma^2(ph)$	2.0000	1.7787
	$\sigma^2(ph)$	2.0000	1.3285	$\sigma^2(pih)$	208.0000	4.8020
	$\sigma^2(ih)$	3.0000	1.4413			
	$\sigma^2(pih)$	144.0000	4.6955			
	$\sigma^2(\delta)^b$	7.8000	0.7197	$\sigma^2(\delta)$	6.2000	0.8974
	$\sigma^2(\Delta)^c$	8.5750	1.1426	$\sigma^2(\Delta)$	6.8750	1.3774
<b>Dichotomous Data</b>						
$n_p = 100$	$\sigma^2(p)$	0.0109	0.0026	$\sigma^2(p)$	0.0108	0.0025
$n_i = 20$	$\sigma^2(i)$	0.0028	0.0018	$\sigma^2(h)$	0.0006	0.0014
$n_h = 2$	$\sigma^2(h)$	0.0007	0.0012	$\sigma^2(ih)$	0.0048	0.0016
	$\sigma^2(pi)$	0.0449	0.0054	$\sigma^2(ph)$	0.0014	0.0018
	$\sigma^2(ph)$	0.0014	0.0016	$\sigma^2(pih)$	0.2323	0.0030
	$\sigma^2(ih)$	0.0021	0.0013			
	$\sigma^2(pih)$	0.1837	0.0057			
	$\sigma^2(\delta)$	0.0076	0.0008	$\sigma^2(\delta)$	0.0065	0.0009
	$\sigma^2(\Delta)$	0.0081	0.0010	$\sigma^2(\Delta)$	0.0070	0.0011
<b>Polytomous Data</b>						
$n_p = 100$	$\sigma^2(p)$	0.3241	0.0524	$\sigma^2(p)$	0.0246	0.0325
$n_i = 20$	$\sigma^2(i)$	0.1270	0.0583	$\sigma^2(h)$	0.1324	0.2499
$n_h = 2$	$\sigma^2(h)$	0.0120	0.0808	$\sigma^2(ih)$	0.4093	0.0829
	$\sigma^2(pi)$	0.3930	0.0179	$\sigma^2(ph)$	0.0651	0.0225
	$\sigma^2(ph)$	0.0140	0.0092	$\sigma^2(pih)$	1.1655	0.0251
	$\sigma^2(ih)$	0.0025	0.0081			
	$\sigma^2(pih)$	0.3170	0.0105			
	$\sigma^2(\delta)$	0.0346	0.0048	$\sigma^2(\delta)$	0.0617	0.0112
	$\sigma^2(\Delta)$	0.0470	0.0393	$\sigma^2(\Delta)$	0.1381	0.1249

<sup>a</sup>Parameter values for dichotomous and polytomous data were based on 5000 trials.

<sup>b</sup> $\sigma^2(\delta)$  refers to relative error variance.

<sup>c</sup> $\sigma^2(\Delta)$  refers to absolute error variance.

Table 2: Estimated Variance Components for the  $p \times i \times h$  Design, Raw Bootstrap Results.

	$\hat{\sigma}^2(p)$	$\hat{\sigma}^2(i)$	$\hat{\sigma}^2(h)$	$\hat{\sigma}^2(pi)$	$\hat{\sigma}^2(ph)$	$\hat{\sigma}^2(ih)$	$\hat{\sigma}^2(pih)$	$\hat{\sigma}^2(\delta)$	$\hat{\sigma}^2(\Delta)$
<b>Normal Data</b>									
T <sup>a</sup>	16.0000	4.0000	1.0000	64.0000	2.0000	3.0000	144.0000	7.8000	8.5750
<i>p</i>	15.8754	4.7047	0.9678	63.6142	2.0007	4.3933	142.5068	7.7438	8.5727
<i>i</i>	19.2233	3.8700	1.0926	61.0402	9.2051	2.8156	136.7320	11.0729	11.8830
<i>h</i>	17.0371	5.5484	0.4736	136.2177	1.0081	1.4792	71.9755	9.1143	9.6656
<i>p, i</i>	19.0250	4.4917	1.1863	60.4233	9.1193	4.1763	135.3943	10.9657	11.8879
<i>p, h</i>	16.8654	6.9290	0.4829	135.0928	0.9967	2.1863	71.0151	9.0284	9.6709
<i>i, h</i>	23.8356	5.2745	0.5427	129.3171	4.6105	1.4039	68.4561	10.4825	11.0527
<i>p, i, h</i>	23.6767	6.5808	0.5897	128.6116	4.5336	2.0691	67.1516	10.3762	11.0518
<b>Dichotomous Data</b>									
T	0.0109	0.0028	0.0007	0.0449	0.0014	0.0021	0.1873	0.0076	0.0081
<i>p</i>	0.0110	0.0032	0.0007	0.0444	0.0013	0.0040	0.1854	0.0075	0.0081
<i>i</i>	0.0133	0.0026	0.0008	0.0426	0.0107	0.0020	0.1778	0.0119	0.0125
<i>h</i>	0.0117	0.0038	0.0003	0.1382	0.0006	0.0011	0.0939	0.0096	0.0100
<i>p, i</i>	0.0132	0.0030	0.0009	0.0422	0.0106	0.0038	0.1761	0.0118	0.0125
<i>p, h</i>	0.0116	0.0052	0.0003	0.1373	0.0006	0.0020	0.0924	0.0095	0.0100
<i>i, h</i>	0.0186	0.0036	0.0004	0.1311	0.0054	0.0010	0.0894	0.0115	0.0119
<i>p, i, h</i>	0.0184	0.0049	0.0004	0.1301	0.0053	0.0019	0.0882	0.0114	0.0119
<b>Polytomous Data</b>									
T	0.3241	0.1270	0.0120	0.3930	0.0140	0.0025	0.3170	0.0346	0.0470
<i>p</i>	0.3197	0.1291	0.0149	0.3902	0.0140	0.0055	0.3138	0.0343	0.0484
<i>i</i>	0.3426	0.1189	0.0149	0.3745	0.0299	0.0022	0.3011	0.0412	0.0547
<i>h</i>	0.3299	0.1263	0.0074	0.5524	0.0071	0.0012	0.1587	0.0351	0.0452
<i>p, i</i>	0.3391	0.1224	0.0152	0.3708	0.0297	0.0052	0.2981	0.0408	0.0547
<i>p, h</i>	0.3266	0.1318	0.0075	0.5469	0.0070	0.0027	0.1573	0.0348	0.0452
<i>i, h</i>	0.3578	0.1200	0.0075	0.5248	0.0150	0.0011	0.1507	0.0375	0.0473
<i>p, i, h</i>	0.3543	0.1253	0.0075	0.5202	0.0148	0.0026	0.1487	0.0371	0.0472

<sup>a</sup>T means parameter values; *p* means bootstrap the *p* facet; others rows are similarly defined.

Table 3: Estimated Variance Components for the  $p \times i \times h$  Design, Bias Adjusted Bootstrap Results.

	$\hat{\sigma}^2(p)$	$\hat{\sigma}^2(i)$	$\hat{\sigma}^2(h)$	$\hat{\sigma}^2(pi)$	$\hat{\sigma}^2(ph)$	$\hat{\sigma}^2(ih)$	$\hat{\sigma}^2(pih)$	$\hat{\sigma}^2(\delta)$	$\hat{\sigma}^2(\Delta)$
<b>Normal Data</b>									
T <sup>a</sup>	<b>16.0000</b>	<b>4.0000</b>	<b>1.0000</b>	<b>64.0000</b>	<b>2.0000</b>	<b>3.0000</b>	<b>144.0000</b>	<b>7.8000</b>	<b>8.5750</b>
<i>p</i>	16.0357	4.0621	0.9476	64.2568	2.0209	2.9539	143.9463	7.8220	8.5727
<i>i</i>	16.0107	4.0737	0.9444	64.2528	2.0087	2.9638	143.9284	7.8152	8.5652
<i>h</i>	16.0290	4.0692	0.9472	64.2422	2.0163	2.9584	143.9511	7.8190	8.5701
<i>p, i</i>	16.0048	4.0856	0.9464	64.2460	2.0135	2.9565	143.9599	7.8180	8.5694
<i>p, h</i>	16.0290	4.0954	0.9457	64.7250	2.0136	2.9379	143.4648	7.8297	8.5807
<i>i, h</i>	16.0218	4.0744	0.9376	64.0642	2.0152	2.9555	144.1182	7.8137	8.5601
<i>p, i, h</i>	16.0690	4.0957	0.9413	65.3481	2.0189	2.9280	142.7999	7.8468	8.5955
<b>Dichotomous Data</b>									
T	<b>0.0109</b>	<b>0.0028</b>	<b>0.0007</b>	<b>0.0449</b>	<b>0.0014</b>	<b>0.0021</b>	<b>0.1873</b>	<b>0.0076</b>	<b>0.0081</b>
<i>p</i>	0.0111	0.0028	0.0007	0.0448	0.0013	0.0021	0.1872	0.0076	0.0081
<i>i</i>	0.0110	0.0027	0.0007	0.0449	0.0013	0.0021	0.1872	0.0076	0.0081
<i>h</i>	0.0110	0.0027	0.0007	0.0442	0.0013	0.0021	0.1879	0.0076	0.0081
<i>p, i</i>	0.0110	0.0028	0.0007	0.0448	0.0013	0.0021	0.1872	0.0076	0.0081
<i>p, h</i>	0.0111	0.0027	0.0007	0.0453	0.0013	0.0021	0.1868	0.0076	0.0081
<i>i, h</i>	0.0110	0.0027	0.0007	0.0438	0.0013	0.0021	0.1883	0.0076	0.0081
<i>p, i, h</i>	0.0110	0.0027	0.0007	0.0446	0.0013	0.0021	0.1875	0.0076	0.0081
<b>Polytomous Data</b>									
T	<b>0.3241</b>	<b>0.1270</b>	<b>0.0120</b>	<b>0.3930</b>	<b>0.0140</b>	<b>0.0025</b>	<b>0.3170</b>	<b>0.0346</b>	<b>0.0470</b>
<i>p</i>	0.3229	0.1251	0.0148	0.3941	0.0141	0.0023	0.3170	0.0347	0.0484
<i>i</i>	0.3229	0.1252	0.0148	0.3942	0.0141	0.0023	0.3170	0.0347	0.0484
<i>h</i>	0.3229	0.1251	0.0148	0.3937	0.0142	0.0023	0.3174	0.0347	0.0484
<i>p, i</i>	0.3228	0.1249	0.0148	0.3942	0.0142	0.0023	0.3169	0.0347	0.0484
<i>p, h</i>	0.3228	0.1251	0.0149	0.3935	0.0142	0.0023	0.3177	0.0347	0.0485
<i>i, h</i>	0.3231	0.1252	0.0150	0.3938	0.0142	0.0023	0.3173	0.0347	0.0485
<i>p, i, h</i>	0.3232	0.1252	0.0147	0.3949	0.0141	0.0023	0.3163	0.0347	0.0483

<sup>a</sup>T means parameter values; *p* means bootstrap the *p* facet; others rows are similarly defined.

Table 4: Estimated Standard Errors for the  $p \times i \times h$  Design, Raw Bootstrap Results

	$SE(p)$	$SE(i)$	$SE(h)$	$SE(pi)$	$SE(ph)$	$SE(ih)$	$SE(pih)$	$SE(\delta)$	$SE(\Delta)$
<b>Normal Data</b>									
T <sup>a</sup>	3.4545	2.3399	1.7580	5.0178	1.3285	1.4413	4.6955	0.7197	1.1426
$p$	3.3992	1.5978	0.6725	4.9511	1.3140	1.2572	4.6280	0.6926	0.7704
$i$	3.0017	2.1973	1.0348	4.7657	2.9621	1.3170	4.9127	1.4142	1.5163
$h$	1.9042	2.0594	0.9345	72.0770	1.2030	1.6387	72.0000	1.3494	1.2298
$p, i$	5.3159	3.0679	1.4190	8.0608	3.7049	2.1199	7.8443	1.8237	1.9705
$p, h$	4.1468	3.2291	1.0572	71.6118	1.5085	2.4717	71.3805	1.4417	1.3668
$i, h$	6.3720	3.1604	1.2205	68.7300	5.0988	1.8133	68.5026	1.2452	1.4870
$p, i, h$	8.0113	4.3141	1.4148	68.5415	5.3098	2.6485	68.0118	1.5184	1.7674
<b>Dichotomous Data</b>									
T	0.0026	0.0018	0.0012	0.0054	0.0016	0.0013	0.0057	0.0008	0.0010
$p$	0.0026	0.0015	0.0005	0.0053	0.0015	0.0014	0.0056	0.0006	0.0007
$i$	0.0027	0.0017	0.0008	0.0050	0.0038	0.0012	0.0060	0.0018	0.0018
$h$	0.0017	0.0016	0.0007	0.0937	0.0010	0.0012	0.0936	0.0021	0.0019
$p, i$	0.0045	0.0026	0.0012	0.0086	0.0047	0.0022	0.0094	0.0022	0.0023
$p, h$	0.0034	0.0029	0.0008	0.0929	0.0015	0.0023	0.0928	0.0021	0.0020
$i, h$	0.0068	0.0024	0.0009	0.0892	0.0060	0.0014	0.0891	0.0014	0.0015
$p, i, h$	0.0079	0.0037	0.0011	0.0884	0.0063	0.0025	0.0883	0.0017	0.0019
<b>Polytomous Data</b>									
T	0.0524	0.0583	0.0808	0.0179	0.0092	0.0081	0.0105	0.0048	0.0393
$p$	0.0512	0.0126	0.0044	0.0175	0.0089	0.0022	0.0104	0.0045	0.0051
$i$	0.0210	0.0526	0.0052	0.0178	0.0069	0.0070	0.0110	0.0033	0.0049
$h$	0.0143	0.0088	0.0435	0.1589	0.0084	0.0043	0.1585	0.0024	0.0218
$p, i$	0.0586	0.0550	0.0076	0.0293	0.0122	0.0075	0.0176	0.0061	0.0077
$p, h$	0.0547	0.0163	0.0436	0.1586	0.0104	0.0052	0.1571	0.0040	0.0221
$i, h$	0.0304	0.0536	0.0436	0.1526	0.0164	0.0064	0.1508	0.0050	0.0231
$p, i, h$	0.0644	0.0565	0.0439	0.1531	0.0177	0.0072	0.1496	0.0063	0.0236

<sup>a</sup>T means parameter values;  $p$  means bootstrap the  $p$  facet; others rows are similarly defined.

Table 5: Estimated Standard Errors for the  $p \times i \times h$  Design, Bias Adjusted Bootstrap Results.

	$SE(p)$	$SE(i)$	$SE(h)$	$SE(pi)$	$SE(ph)$	$SE(ih)$	$SE(pih)$	$SE(\delta)$	$SE(\Delta)$
<b>Normal Data</b>									
T <sup>a</sup>	<b>3.4545</b>	<b>2.3399</b>	<b>1.7580</b>	<b>5.0178</b>	<b>1.3285</b>	<b>1.4413</b>	<b>4.6955</b>	<b>0.7197</b>	<b>1.1426</b>
$p$	<i>3.4336</i>	1.6004	0.6727	<i>5.0011</i>	<i>1.3272</i>	1.2598	<i>4.6748</i>	0.6996	0.7704
$i$	3.0475	<i>2.3129</i>	1.0407	5.0165	3.0800	<i>1.3863</i>	5.1712	1.4687	1.5698
$h$	2.8225	3.5056	<i>1.8690</i>	144.0396	2.4061	3.2773	144.0001	2.6836	2.4413
$p, i$	5.4185	3.2320	1.4286	8.5708	3.8657	2.2349	8.3406	1.9017	2.0291
$p, h$	4.8605	4.1688	2.1055	144.2768	3.0475	3.7240	144.2031	2.8624	2.6761
$i, h$	5.4463	4.5811	2.3857	144.2592	4.9791	3.8174	144.2158	3.4665	3.3758
$p, i, h$	7.7691	5.5599	2.7501	144.7578	5.9670	4.5293	144.6290	3.8539	3.8263
<b>Dichotomous Data</b>									
T	<b>0.0026</b>	<b>0.0018</b>	<b>0.0012</b>	<b>0.0054</b>	<b>0.0016</b>	<b>0.0013</b>	<b>0.0057</b>	<b>0.0008</b>	<b>0.0010</b>
$p$	<i>0.0026</i>	0.0015	0.0005	<i>0.0054</i>	<i>0.0015</i>	0.0014	<i>0.0056</i>	0.0007	0.0007
$i$	0.0028	<i>0.0018</i>	0.0008	0.0053	0.0039	<i>0.0012</i>	0.0063	0.0019	0.0019
$h$	0.0025	0.0027	<i>0.0014</i>	0.1873	0.0020	0.0024	0.1873	0.0041	0.0039
$pi$	0.0047	0.0027	0.0013	0.0092	0.0049	0.0023	0.0100	0.0024	0.0024
$ph$	0.0042	0.0035	0.0017	0.1875	0.0030	0.0032	0.1875	0.0043	0.0041
$ih$	0.0057	0.0036	0.0019	0.1876	0.0059	0.0030	0.1876	0.0050	0.0048
$pih$	0.0077	0.0047	0.0023	0.1878	0.0072	0.0041	0.1878	0.0054	0.0053
<b>Polytomous Data</b>									
T	<b>0.0524</b>	<b>0.0583</b>	<b>0.0808</b>	<b>0.0179</b>	<b>0.0092</b>	<b>0.0081</b>	<b>0.0105</b>	<b>0.0048</b>	<b>0.0393</b>
$p$	<i>0.0517</i>	0.0126	0.0044	<i>0.0177</i>	<i>0.0090</i>	0.0022	<i>0.0105</i>	0.0046	0.0051
$i$	0.0211	<i>0.0553</i>	0.0051	0.0187	0.0072	<i>0.0074</i>	0.0116	0.0034	0.0050
$h$	0.0204	0.0116	<i>0.0869</i>	0.3172	0.0168	0.0087	0.3171	0.0047	0.0436
$p, i$	0.0593	0.0579	0.0075	0.0312	0.0125	0.0079	0.0187	0.0063	0.0078
$p, h$	0.0576	0.0179	0.0872	0.3181	0.0211	0.0092	0.3175	0.0080	0.0442
$i, h$	0.0312	0.0573	0.0872	0.3184	0.0197	0.0134	0.3176	0.0069	0.0441
$p, i, h$	0.0661	0.0603	0.0878	0.3196	0.0243	0.0144	0.3182	0.0101	0.0449

<sup>a</sup>T means parameter values;  $p$  means bootstrap the  $p$  facet; others rows are similarly defined.

Table 6: Estimated Variance Components for the  $p \times (i : h)$  Design, Raw Bootstrap Results.

	$\hat{\sigma}^2(p)$	$\hat{\sigma}^2(h)$	$\hat{\sigma}^2(i:h)$	$\hat{\sigma}^2(ph)$	$\hat{\sigma}^2(pi:h)$	$\hat{\sigma}^2(\delta)$	$\hat{\sigma}^2(\Delta)$
<b>Normal Data</b>							
T <sup>a</sup>	16.0000	1.0000	7.0000	2.0000	208.0000	6.2000	6.8750
<i>p</i>	15.9737	0.9892	9.2269	1.9262	208.0369	6.1025	6.8212
<i>h</i>	22.2862	0.3163	7.1592	-4.2120	210.1171	3.0853	3.4159
<i>i, h</i>	22.1291	0.6625	6.7312	6.1504	197.6381	8.0162	8.5157
<i>p, h</i>	22.0726	0.2714	9.2417	-4.1702	208.0174	3.0538	3.4139
<i>p, i, h</i>	22.1855	0.7237	8.7879	6.0087	197.7164	7.9376	8.5143
<b>Dichotomous Data</b>							
T	0.0108	0.0006	0.0048	0.0014	0.2323	0.0065	0.0070
<i>p</i>	0.0109	0.0007	0.0072	0.0014	0.2321	0.0064	0.0069
<i>h</i>	0.0174	0.0002	0.0049	-0.0051	0.2344	0.0032	0.0035
<i>i, h</i>	0.0173	0.0004	0.0046	0.0065	0.2205	0.0087	0.0091
<i>p, h</i>	0.0173	0.0002	0.0072	-0.0051	0.2321	0.0032	0.0035
<i>p, i, h</i>	0.0174	0.0005	0.0069	0.0063	0.2206	0.0087	0.0091
<b>Polytomous Data</b>							
T	0.2046	0.1324	0.4093	0.0651	1.1655	0.0617	0.1381
<i>p</i>	0.2057	0.1282	0.4241	0.0645	1.1656	0.0608	0.1347
<i>h</i>	0.2688	0.0551	0.4128	0.0040	1.1772	0.0308	0.0679
<i>i, h</i>	0.2671	0.0736	0.3882	0.0613	1.1075	0.0583	0.1048
<i>p, h</i>	0.2674	0.0541	0.4241	0.0032	1.1655	0.0301	0.0670
<i>p, i, h</i>	0.2664	0.0747	0.4031	0.0612	1.1081	0.0580	0.1048

<sup>a</sup>T means parameter values; *p* means bootstrap the *p* facet; others rows are similarly defined.

Table 7: Estimated Variance Components for the  $p \times (i:h)$  Design, Biased Adjusted Bootstrap Results.

	$\hat{\sigma}^2(p)$	$\hat{\sigma}^2(h)$	$\hat{\sigma}^2(i:h)$	$\hat{\sigma}^2(ph)$	$\hat{\sigma}^2(pi:h)$	$\hat{\sigma}^2(\delta)$	$\hat{\sigma}^2(\Delta)$
<b>Normal Data</b>							
T <sup>a</sup>	<b>16.0000</b>	<b>1.0000</b>	<b>7.0000</b>	<b>2.0000</b>	<b>208.0000</b>	<b>6.2000</b>	<b>6.8750</b>
<i>p</i>	<i>15.9738</i>	0.9603	7.0756	<i>1.9264</i>	<i>208.0393</i>	6.1642	6.8212
<i>h</i>	26.3577	<i>0.6135</i>	7.0884	-8.4625	208.0367	0.9696	1.4536
<i>i, h</i>	16.9281	0.9708	<i>7.0855</i>	1.8989	208.0401	6.1504	6.8130
<i>p, h</i>	26.3560	0.6080	7.0907	-8.4630	208.0171	0.9689	1.4502
<i>p, i, h</i>	15.9884	0.9573	7.0957	1.9092	208.0130	6.1549	6.8110
<b>Dichotomous Data</b>							
T	<b>0.0108</b>	<b>0.0006</b>	<b>0.0048</b>	<b>0.0014</b>	<b>0.2323</b>	<b>0.0065</b>	<b>0.0070</b>
<i>p</i>	<i>0.0109</i>	0.0006	0.0048	<i>0.0014</i>	<i>0.2321</i>	0.0065	0.0069
<i>h</i>	0.0224	<i>0.0004</i>	0.0048	-0.0102	0.2321	0.0007	0.0010
<i>i, h</i>	0.0115	0.0007	<i>0.0048</i>	0.0013	0.2321	0.0065	0.0069
<i>p, h</i>	0.0224	0.0004	0.0048	-0.0102	0.2321	0.0007	0.0010
<i>p, i, h</i>	0.0108	0.0006	0.0049	0.0014	0.2321	0.0065	0.0069
<b>Polytomous Data</b>							
T	<b>0.2046</b>	<b>0.1324</b>	<b>0.4093</b>	<b>0.0651</b>	<b>1.1655</b>	<b>0.0617</b>	<b>0.1381</b>
<i>p</i>	<i>0.2057</i>	0.1263	0.4084	<i>0.0645</i>	<i>1.1656</i>	0.0614	0.1347
<i>h</i>	0.2633	<i>0.1076</i>	0.4087	0.0068	1.1656	0.0325	0.0965
<i>i, h</i>	0.2379	0.1268	<i>0.4086</i>	0.0643	1.1657	0.0613	0.1349
<i>p, h</i>	0.2651	0.1056	0.4084	0.0052	1.1655	0.0317	0.0947
<i>p, i, h</i>	0.2047	0.1256	0.4084	0.0653	1.1658	0.0618	0.1348

<sup>a</sup>T means parameter values; *p* means bootstrap the *p* facet; others rows are similarly defined.

Table 8: Estimated Standard Errors for the  $p \times (i:h)$  Design, Raw Bootstrap Results.

	$SE(p)$	$SE(h)$	$SE(i:h)$	$SE(ph)$	$SE(pi:h)$	$SE\delta$	$SE(\Delta)$
<b>Normal Data</b>							
T <sup>a</sup>	3.2761	2.0872	2.0836	1.7787	4.8020	0.8974	1.3774
$p$	3.5468	0.9547	1.5634	2.3120	5.2771	0.8610	0.9543
$h$	6.4519	1.1976	1.5102	6.2285	4.0544	3.1118	3.4731
$i, h$	7.1032	1.7564	2.4183	7.1428	6.0371	3.5523	3.9387
$p, h$	7.3496	1.4106	2.4038	6.4575	7.0312	3.1429	3.5389
$p, i, h$	8.4163	2.0716	3.3170	7.5880	9.5156	3.7208	4.1542
<b>Dichotomous Data</b>							
T	0.0025	0.0014	0.0016	0.0018	0.0030	0.0009	0.0011
$p$	0.0029	0.0007	0.0015	0.0025	0.0037	0.0009	0.0009
$h$	0.0067	0.0008	0.0011	0.0065	0.0030	0.0033	0.0035
$i, h$	0.0073	0.0012	0.0019	0.0076	0.0042	0.0038	0.0040
$p, h$	0.0072	0.0010	0.0021	0.0068	0.0045	0.0033	0.0036
$p, i, h$	0.0083	0.0016	0.0028	0.0081	0.0063	0.0040	0.0042
<b>Polytomous Data</b>							
T	0.0325	0.2499	0.0829	0.0225	0.0251	0.0112	0.1249
$p$	0.0350	0.0428	0.0255	0.0262	0.0274	0.0111	0.0175
$h$	0.0658	0.1407	0.0592	0.0624	0.0207	0.0312	0.0903
$i, h$	0.0692	0.1635	0.0966	0.0659	0.0311	0.0329	0.1003
$p, h$	0.0762	0.1462	0.0664	0.0650	0.0364	0.0318	0.0912
$p, i, h$	0.0826	0.1673	0.1025	0.0697	0.0491	0.0342	0.1010

<sup>a</sup>T means parameter values;  $p$  means bootstrap the  $p$  facet; others rows are similarly defined.

Table 9: Estimated Standard Errors for the  $p \times (i:h)$  Design, Biased Adjusted Bootstrap Results.

	$SE(p)$	$SE(h)$	$SE(i:h)$	$SE(ph)$	$SE(pi:h)$	$SE\delta$	$SE(\Delta)$
<b>Normal Data</b>							
T <sup>a</sup>	<b>3.2761</b>	<b>2.0872</b>	<b>2.0836</b>	<b>1.7787</b>	<b>4.8020</b>	<b>0.8974</b>	<b>1.3774</b>
$p$	<i>3.2465</i>	0.8625	1.5246	<i>1.7561</i>	<i>4.7915</i>	0.8697	0.9543
$h$	12.5606	<i>2.3885</i>	1.5006	12.4519	3.4397	6.2242	6.9467
$i, h$	7.1060	3.5220	<i>2.5456</i>	14.3357	6.3549	7.1428	7.9148
$p, h$	13.1861	2.6681	2.3922	12.7164	6.7500	6.3515	7.0802
$p, i, h$	14.8439	4.0766	3.4848	15.1720	9.8439	7.5546	8.3458
<b>Dichotomous Data</b>							
T	<b>0.0025</b>	<b>0.0014</b>	<b>0.0016</b>	<b>0.0018</b>	<b>0.0030</b>	<b>0.0009</b>	<b>0.0011</b>
$p$	<i>0.0025</i>	0.0007	0.0014	<i>0.0019</i>	<i>0.0029</i>	0.0009	0.0009
$h$	0.0132	<i>0.0017</i>	0.0011	0.0131	0.0019	0.0065	0.0070
$i, h$	0.0073	0.0026	<i>0.0020</i>	0.0153	0.0044	0.0076	0.0081
$p, h$	0.0136	0.0019	0.0021	0.0134	0.0038	0.0067	0.0072
$p, i, h$	0.0154	0.0032	0.0030	0.0162	0.0062	0.0081	0.0086
<b>Polytomous Data</b>							
T	<b>0.0325</b>	<b>0.2499</b>	<b>0.0829</b>	<b>0.0225</b>	<b>0.0251</b>	<b>0.0112</b>	<b>0.1249</b>
$p$	<i>0.0322</i>	0.0270	0.0227	<i>0.0224</i>	<i>0.0244</i>	0.0112	0.0175
$h$	0.1264	<i>0.2813</i>	0.0590	0.1248	0.0169	0.0624	0.1806
$i, h$	0.0692	0.3272	<i>0.1017</i>	0.1319	0.0327	0.0659	0.2007
$p, h$	0.1338	0.2842	0.0655	0.1286	0.0346	0.0643	0.1824
$p, i, h$	0.1416	0.3293	0.1076	0.1387	0.0505	0.0692	0.2021

<sup>a</sup>T means parameter values;  $p$  means bootstrap the  $p$  facet; others rows are similarly defined.