Monte Carlo Approach to Model Fit Evaluation in Structural Equation Modeling: How to Specify Trivial Misspecification

Sunthud Pornprasertmanit

Wei Wu

Todd D. Little

University of Kansas

Author Note

We appreciate help from the members of the KU Model Evaluation Workgroup, as well as all members of the Center for Research Methods and Data Analysis, University of Kansas, who provided feedback and helpful suggestions for this project. Partial support for this project was provided by grant NSF 1053160 (Wei Wu & Todd D. Little, co-PIs) and the Center for Research Methods and Data Analysis at the University of Kansas (Todd D. Little, director). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

Contact: Sunthud Pornprasertmanit, Department of Psychology and Center for Research Method and Data Analysis, University of Kansas. Email: psunthud@ku.edu

Abstract

This presentation compares three methods of specifying trivial misspecification in Monte Carlo approach to model fit evaluation in structural equation modeling: fixed, random, and maximal, in terms of their rejection rate of the trivially misspecified and severely misspecified models. The simulation study shows that the fixed method leads to higher rejection rate of both trivially and severely misspecified models than the random and maximal methods. By using the random or maximal methods, as sample size increases, the rejection rate of the trivially misspecified model can be controlled to be low while allowing the rejection rate of the severely misspecified model to be high.

Monte Carlo Approach to Model Fit Evaluation in Structural Equation Modeling:

How to Specify Trivial Misspecification

An important issue in structural equation modeling is to evaluate whether a hypothesized model fits an obtained data. Traditionally, model fit is assessed by the likelihood ratio test (LR) and cutoffs associated with the practical fit indices (e.g., RMSEA). It is well known that LR is a test of exact model fit which is not realistic in practice. In addition, LR is sensitive to trivial model misspecifications if sample size is large. Although some of the practical fit indices attempts to measure approximation model fit, the cutoffs proposed for the practical fit indices are not applicable to all hypothesized models (Marsh et al., 2004; Wu, West, & Taylor, 2009). These problems lead to misleading inference about the model adequacy. To solve the problems, Millsap (2007, 2010, in press; Millsap & Lee, 2008, 2009) suggested a Monte Carlo (also known as parametric bootstrap) approach which can accommodate trivial misspecification in the model fit evaluation as well as adjust the cutoff criterion of a fit index according to the nature of any hypothesized model.

To derive the cutoff criterion for a fit index, the Monte Carlo approach involves three steps: 1) simulate a large number of datasets from a population which is equal to the hypothesized model plus some trivial misspecification(s). In doing so, the hypothesized model is a good approximation of the population but not exactly equal to the population, 2) fit the hypothesized model to each generated dataset and record the fit index, and 3) the saved fit indices then form an empirical sampling distribution of the fit index. The cutoff criterion is the critical value in the empirical sampling distribution given a priori type I error rate (e.g., .05).

How to correctly specify trivial misspecification in the first step of the Monte Carlo approach is a challenging issue. Millsap (2010) proposed using an exemplar of maximally acceptable misspecifications (e.g., a misspecified cross loading of size .3). We refer to this method of specifying trivial misspecification as the *fixed method*. For example, Figure 1a shows a hypothesized model without any misspecifications. The empirical sampling distribution of RMSEA when the population is exactly equal to the hypothesized model is shown in Figure 2a. Based on the distribution, we can derive a cutoff of .032 for RMSEA. If the trivial misspecification is added in the population as shown in Figure 1b, the empirical sampling distribution (see Figure 2b) suggests that the RMSEA cutoff is .086. A disadvantage of this approach is that it does not account for other possible misspecifications. For this reason, we propose two new and more realistic methods for setting trivial misspecifications: random and maximal methods. The *random method* treats a model misspecification as random and it has a distribution. Thus, the random method can account for a wide range of possible misspecified models. For example, all possible cross-loadings shown as the orange lines in Figure 1c are drawn from a uniform distribution ranged from -0.3 to 0.3 such that the datasets across all replications are built from different sets of cross-loadings. The empirical sampling distribution based on the random approach is shown in Figure 2c which led to a cutoff of .155 for RMSEA. The *maximal method* also accommodates the fact that there could be a range of trivial misspecifications. However, instead of using a random value within the range, the maximal method selects the misspecification that provides maximum misfit and uses it to define the population. For example, the program routine will find the combination of cross-loadings in Figure 1c that provides the maximum misfit, which is defined in terms of population RMSEA. The empirical sampling distribution based on the maximal approach is shown in Figure 2d with a cutoff of .178 for RMSEA.

Simulation Study

A simulation study is conducted to evaluate the performance of the three approaches to specify trivial misspecifications. We examined several commonly reported fit indices including RMSEA, CFI, TLI, and SRMR. The hypothesized model is a three-factor confirmatory factor model with three indicators for each factor. Factor correlations were 0.5 and standardized factor loadings were 0.7. The data generation (population)model is the hypothesized model plus two cross-loadings (Figure 1b). The amount of cross-loadings of 0.3 is the limit of the size of trivial misspecification—researchers may increase or decrease the degree of trivial cross-loadings based on their own definition of trivial misspecification.

Conditions

1. The amount of misspecification in the tested model represented by the size of the two cross-loadings: 0 (null), 0.3 (trivial), and 0.7 (severe).
2. Sample sizes: 125, 250, 500, and 1000
3. Ways to specify trivial misspecifications in the Monte Carlo approach: a) none, b) fixed with the correct cross-loadings at 0.3 (from Factor 2 to Indicator 1 and from Factor 3 to Indicator 2) f), c) fixed with incorrect cross loadings at 0.3 (from Factor 1 to Indicator 6 and from Factor 2 to Indicator 9), d) random method that draws the value of the standardized cross-loading from a uniform distribution ranging from -0.3 to 0.3, e) random method that draws the value of the standardized cross-loading from a normal distribution with the mean of 0 and the standard deviation of 0.1, and f) the maximal method with the possible range of standardized cross-loadings was -0.3 to 0.3.

Each condition was analyzed using 1,000 replications. The simulation was run by the simsem package in R (Pornprasertmanit, Miller, & Schoemann, 2012). The example code is available at https://github.com/simsem/simsem/wiki

Data Analysis

The performance of the methods is evaluated in terms of the rejection rate of trivially and severely misspecification models. Similar to type I error rate, the rejection rate should be close to 0 if the amount of misfit in the population is none or trivial. This concept is similar to type I error rate. In contrast, the rejection rate should be high (e.g., .80) if the amount of misfit in the population is severe or unacceptable. This concept is similar to statistical power.

Results

Figure 3 shows the rejection rate given six types of misspecification. The black solid, red dashed, and green dotted lines represent the result from the hypothesized model with null, trivial, and severe misspecifications respectively. We will show the results using RMSEA only because other fit indices provided similar results.

Null misspecification. The rejection rate in this condition should be close to 0. The results from all types of misspecifications did not provide the rejection rate over than .05. Without any misspecification, the rejection rate was .05, which is equal to the specified alpha level.

Trivial misspecification. The proportion of rejection in this condition should be close to 0 as well. If the trivial misspecification was not added in the simulated data, the rejection rate was high which is undesirable. The fixed method led to an increasing rejection rate when the sample size increased. In contrast, the random and maximal methods offered rejection rates that were close to 0.

Severe misspecification. The rejection rate should be close to 1. The rejection rate of the severely misspecified model when the trivial misspecification was not added in the simulated data was close to 1. The fixed method also yielded the similar result. In comparision, the random and maximal method yielded low rejection rate of the severely misspecified model, especially when sample size is low (*N* = 125). As sample size increased, rejection rate also increased.

We also examined the amount of population RMSEA provided by the different methods of specifying trivial misspecification. The population RMSEA is 0 when there is no misspecification. The population RMSEAs of the model with the fixed, random, and maximal methods of trivial misspecification are .092, .159, and .213, respectively. From the results shown above, the rejection rate depends on the population RMSEA. When the methods of specifying trivial misspecification provided low population RMSEA (none or fixed method), the hypothesized model with trivial and severe misspecifications is rejected. When the methods of specifying trivial misspecification provided high population RMSEA (random or maximal method), the hypothesized model with severe misspecification is rejected.

Discussion and Conclusion

The main purpose of this presentation is to evaluate three methods of specifying trivial misspecification in Monte Carlo approach to model fit evaluation. When exact model fit is considered, the hypothesized model with trivial misspecification are mostly rejected. The fixed method lead to overrejection of the hypothesized model with trivial misfit, especially in high sample size; however, it lead to high power detect severely misspecified model. On the other hand, the random and maximal methods were more likely to retain a trivially misspecified model. However, they are not sensitive to severe misfit, especially with small sample size.

The population RMSEA which measures the population misfit can explain the differential performance of the three approaches to specify trivial misspecification (none < fixed < random < maximal). Loosely speaking, the population misfit is related to the type I error and statistical power. If the method of specifying trivial misspecification contains greater misfit, the rejection rate of the hypothesized models with trivial and severe misfits are lower. The rejection rates of the hypothesized model with trivial or severe misfits are higher when the sample size is higher. The random and maximal methods with high sample size not only provide enough rejection rates for the severely misspecified model but also control the rejection rates for the trivial misspecified model. Therefore, we recommend a priori statistical power analysis for researchers to pick the method of specifying trivial misspecification (random and maximal methods in this study) and enough sample size such that the rejection rate of the severe misspecified model is high and the rejection rate of the trivial misspecified model is low.

References

Marsh, H. W., Hau, K. T., & Wen Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling*, *11*, 320-324.

Millsap, R. (2007). Structural equation modeling made difficult. *Personality and Individual Differences*, *42*, 875-881.

Millsap, R. E. (September, 2010). *A simulation paradigm for evaluating "approximate fit" in latent variable modeling*. Paper presented at the Conference Honoring the Scientific Contributions of Michael W. Browne", Ohio State University, Columbus, OH.

Millsap, R.E. (in press). A simulation paradigm for evaluating model fit. In M. Edwards and R. MacCallum (Eds.) *Current Issues in the Theory and Application of Latent Variable Models*. New York: Routledge.

Millsap, R.E. & Lee, S (2008). *Approximate fit in SEM without a priori cutpoints*. Paper presented at the Annual meeting of Society of Multivariate Experimental Psychology.

Millsap, R.E. & Lee, S (2009). *Approximate fit in SEM without cutpoints*. Paper presented at the Psychometric Society, Annual Meeting, Cambridge University, Cambridge, England.

Pornprasertmanit, S., Miller, P. J., & Schoemann, A. (2012). *simsem: Simulated Structural Equation Modeling* [Computer Software]. Available from https://github.com/simsem/simsem

Wei, W., West, S. G., Taylor, A. B. (2009). Evaluating model fit for growth curve models: Integration of fit indices from SEM and MLM frameworks. *Psychological Methods, 14*, 183-201.

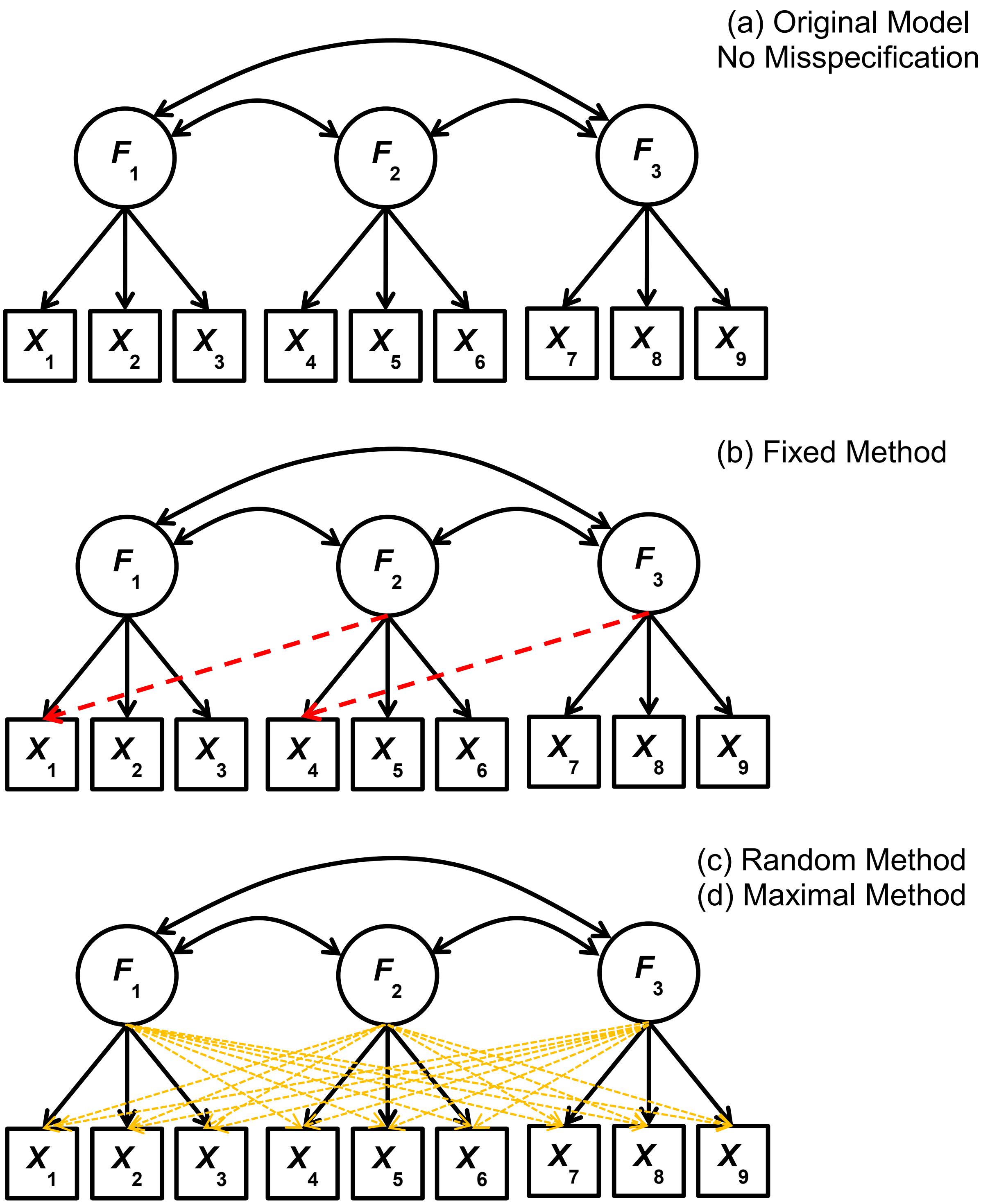

Figure 1. The methods to specify trivial misspecification

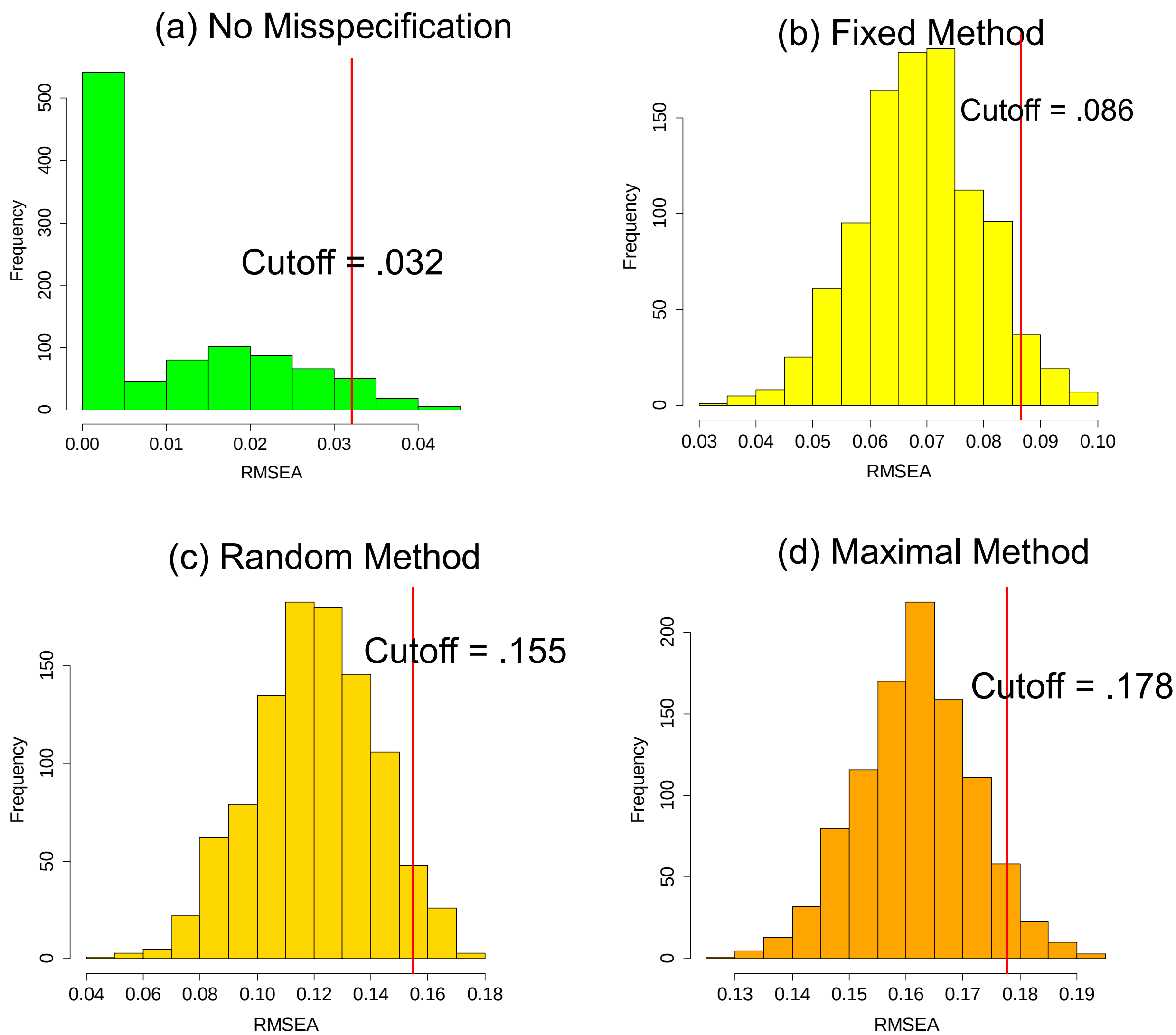

Figure 2. The empirical sampling distribution based on different methods to specify trivial misspecification

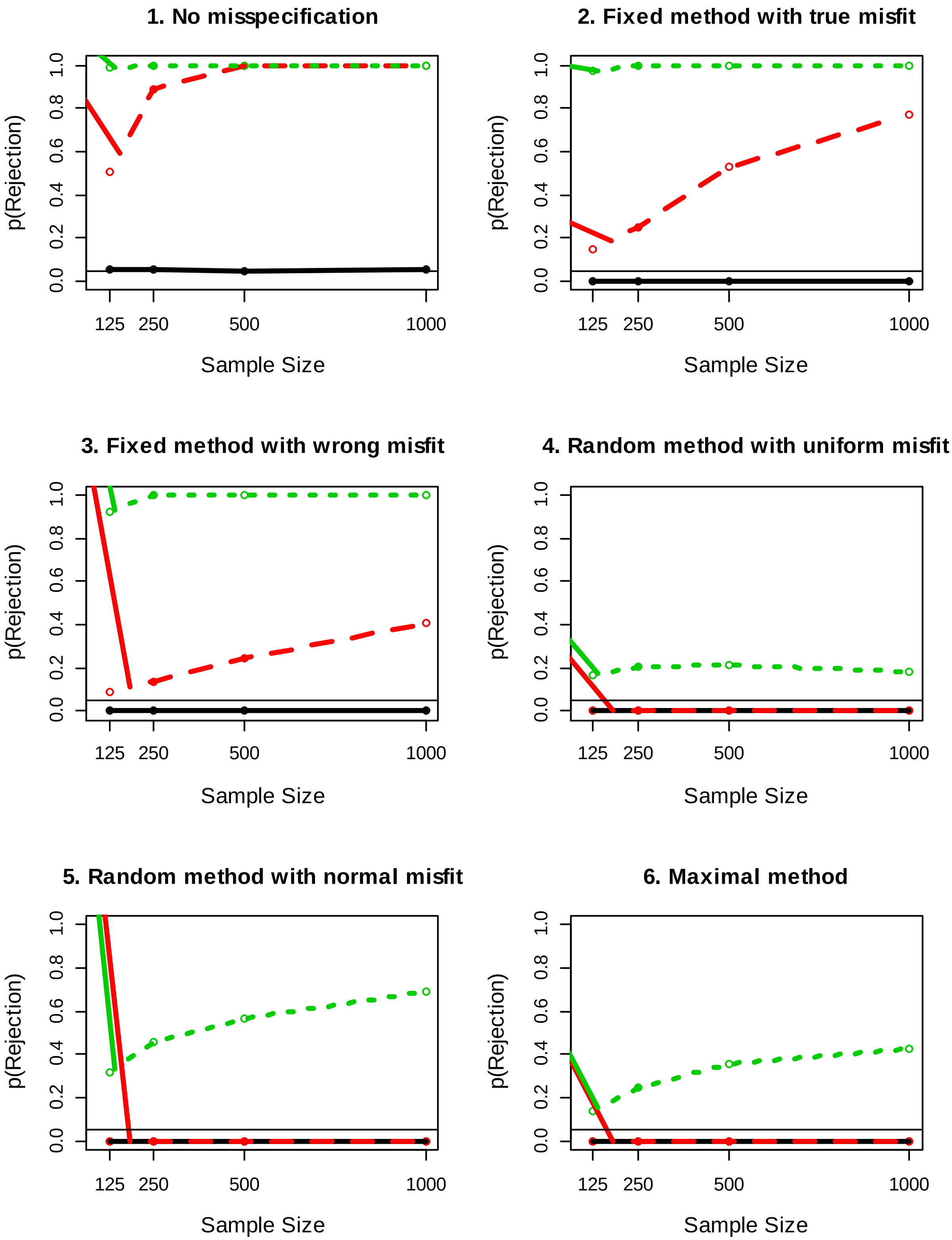

Figure 3. The rejection rate of each condition