

# Supplementary Material for “Hypothesis Testing for Hierarchical Structures in Cognitive Diagnosis Models”

CHENCHEN MA<sup>1</sup> AND GONGJUN XU<sup>\*1</sup>

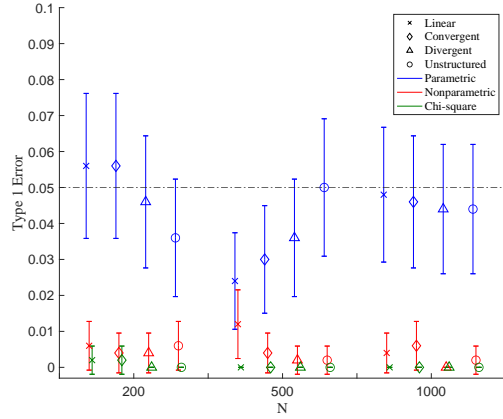
<sup>1</sup>*DEPARTMENT OF STATISTICS, UNIVERSITY OF MICHIGAN, WEST HALL, 1085 S UNIVERSITY AVE,  
ANN ARBOR, MI 48109, USA*

This supplementary material includes Type 1 Errors in Section 4.2, the  $Q$ -matrix for ECPE data in Section 5 and additional simulation results. Specifically, bootstrap results for DINA and GDINA models under both null hypothesis and alternative hypothesis with different sample sizes and noise levels are presented in the following Figures 1–4. Codes for simulations and real data analysis are available in a separate supplementary material.

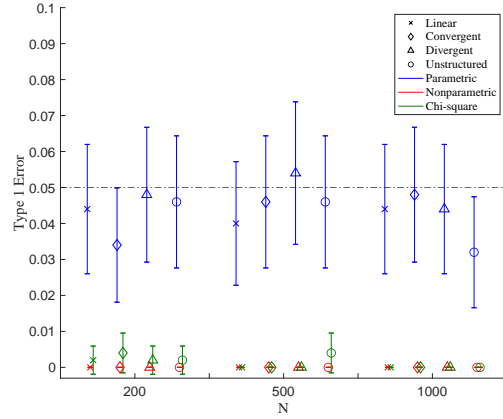
---

<sup>\*</sup>Corresponding author. Email: gongjun@umich.edu

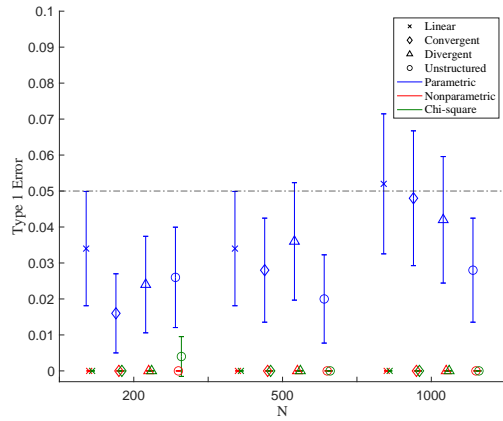
# 1 Type 1 Errors



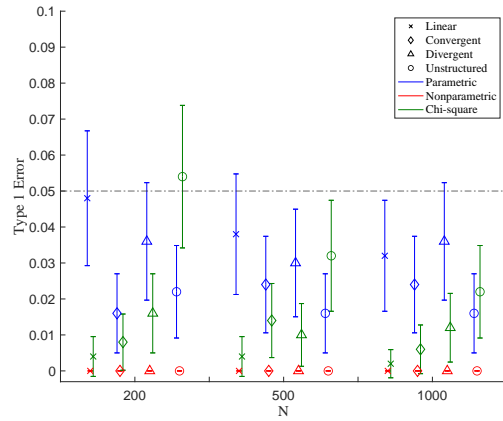
(a) DINA,  $\theta_j^+ = 0.9, \theta_j^- = 0.1$



(b) DINA,  $\theta_j^+ = 0.8, \theta_j^- = 0.2$



(c) GDINA,  $\theta_j^+ = 0.9, \theta_j^- = 0.1$



(d) GDINA,  $\theta_j^+ = 0.8, \theta_j^- = 0.2$

Figure 1: Type 1 Errors. Different colors indicate different testing procedures. Different marker shapes stand for different hierarchical structures. The middle points are the means of the type I errors and the vertical errors bars with  $\pm 2$  s.e. are constructed based on 500 replications.  $\theta_j^+ = 0.9, \theta_j^- = 0.1$  corresponds to the case with low noises, and  $\theta_j^+ = 0.8, \theta_j^- = 0.2$  corresponds to the case with high noises.

## 2 $Q$ -matrix for ECPE data

Item	Attributes		
	Mor.rules ( $\alpha_1$ )	Coh.rules ( $\alpha_2$ )	Lex.rules ( $\alpha_3$ )
1	1	1	0
2	0	1	0
3	1	0	1
4	0	0	1
5	0	0	1
6	0	0	1
7	1	0	1
8	0	1	0
9	0	0	1
10	1	0	0
11	1	0	1
12	1	0	1
13	1	0	0
14	1	0	0
15	0	0	1
16	1	0	1
17	0	1	1
18	0	0	1
19	0	0	1
20	1	0	1
21	1	0	1
22	0	0	1
23	0	1	0
24	0	1	0
25	1	0	0
26	0	0	1
27	1	0	0
28	0	0	1

Table 1: The  $Q$ -matrix for ECPE data. "Mor." is short for "morphosyntactic", "Coh." is short for "cohesive", and "Lex." is short for "lexical".

### 3 Bootstrap Results under the DINA model

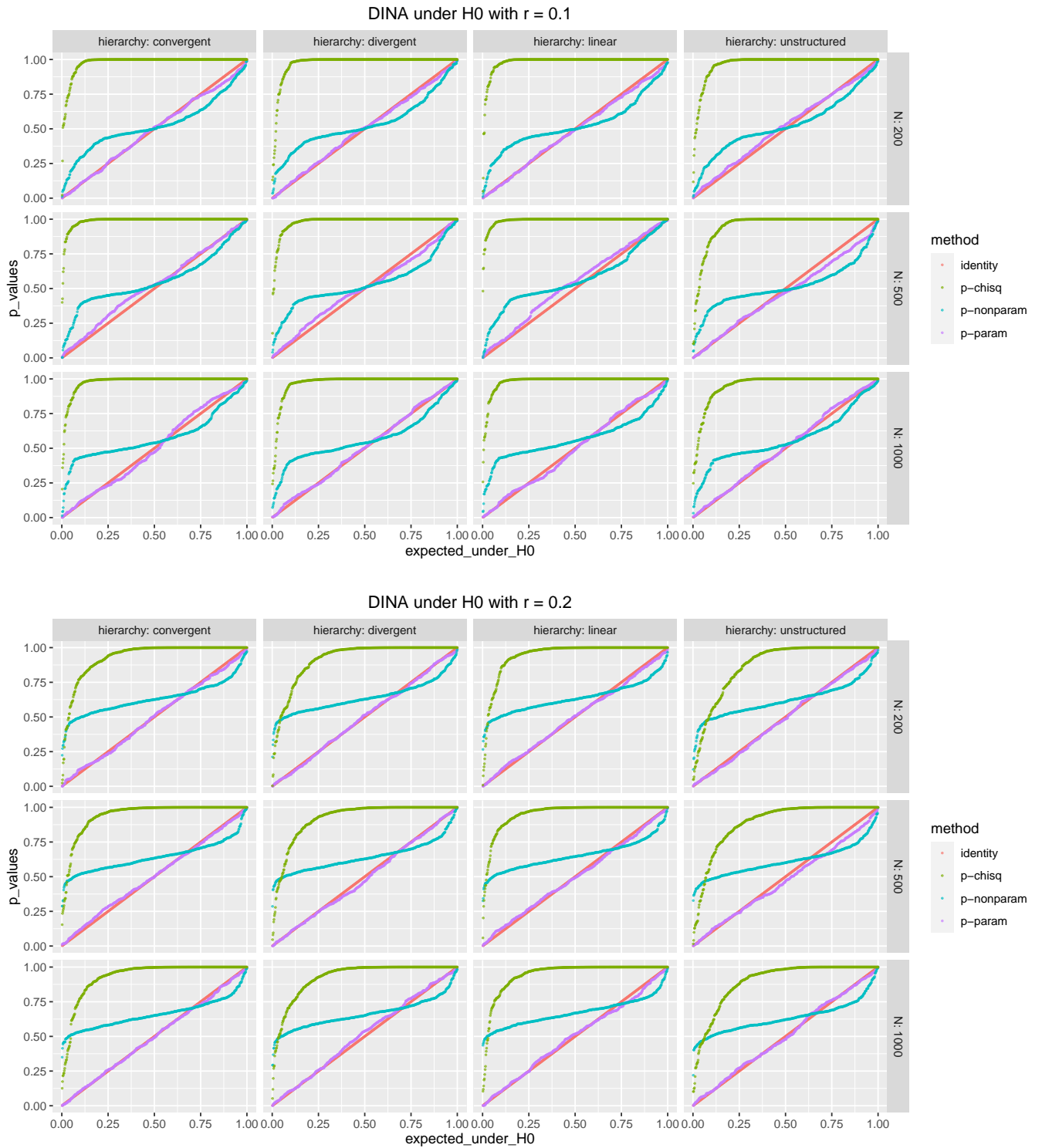


Figure 2: Bootstrap results for DINA under null hypothesis

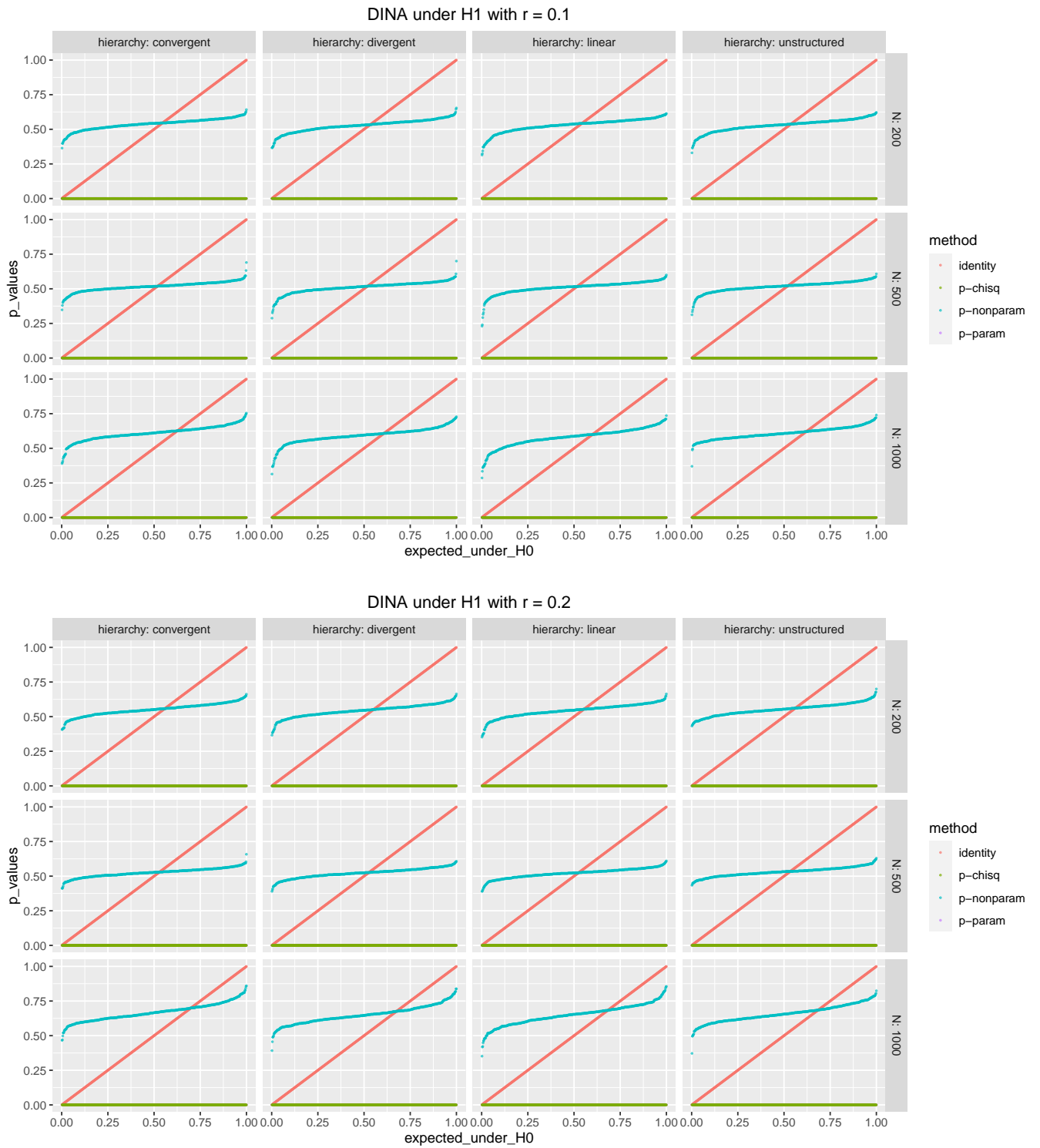


Figure 3: Bootstrap results for DINA under alternative hypothesis

## 4 Bootstrap Results under the GDINA model

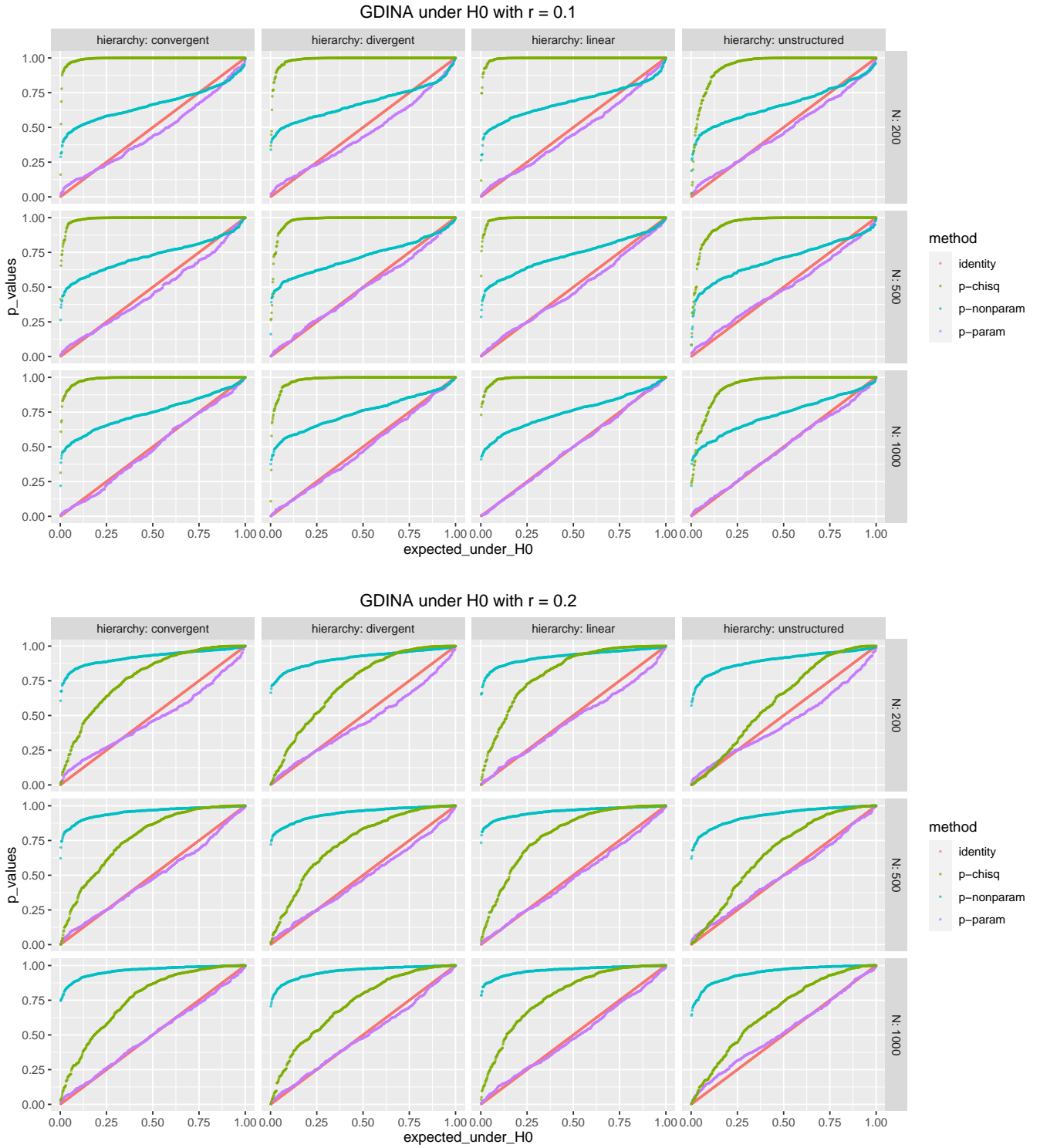


Figure 4: Bootstrap results for GDINA under null hypothesis

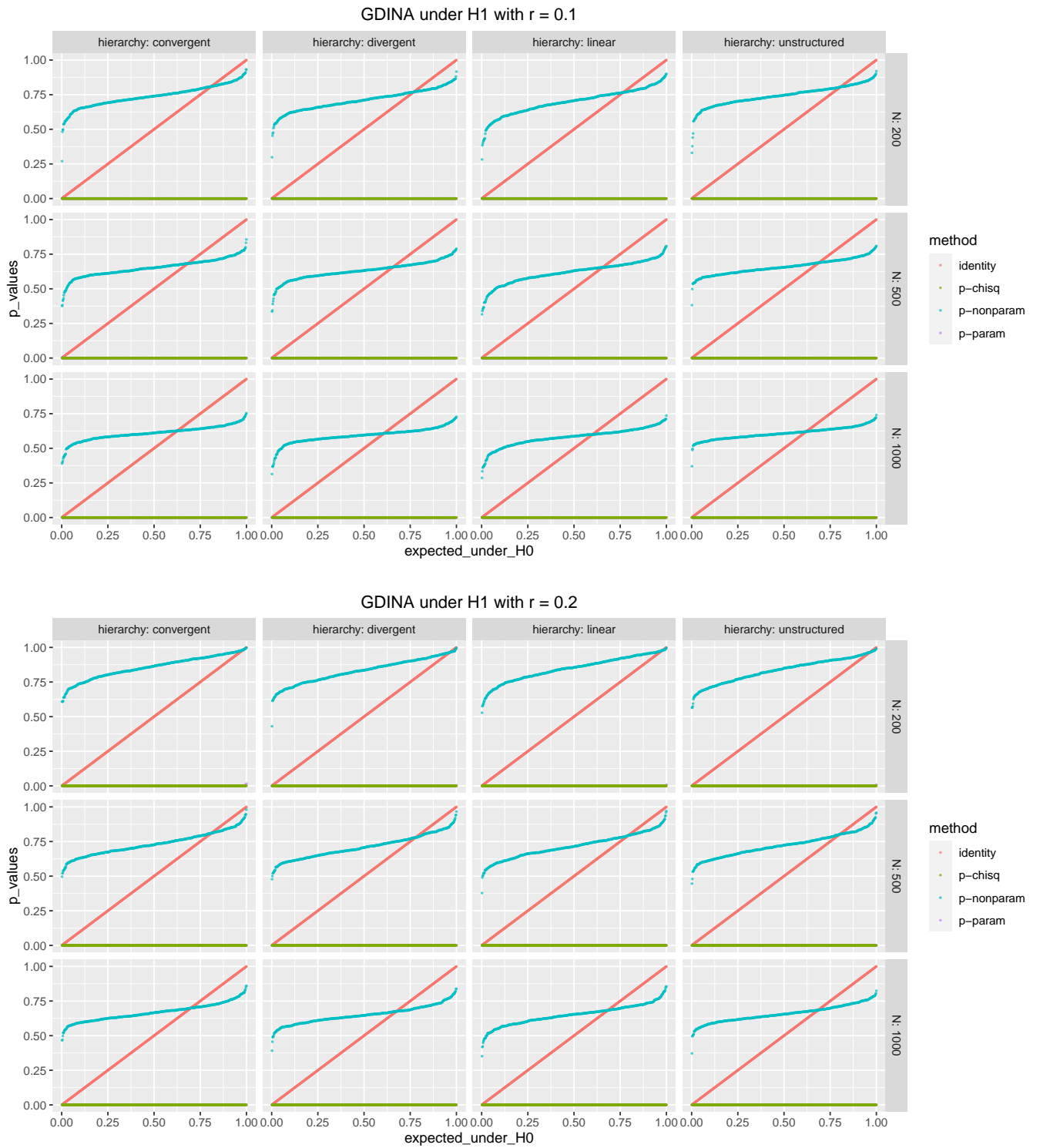


Figure 5: Bootstrap results for GDINA under alternative hypothesis