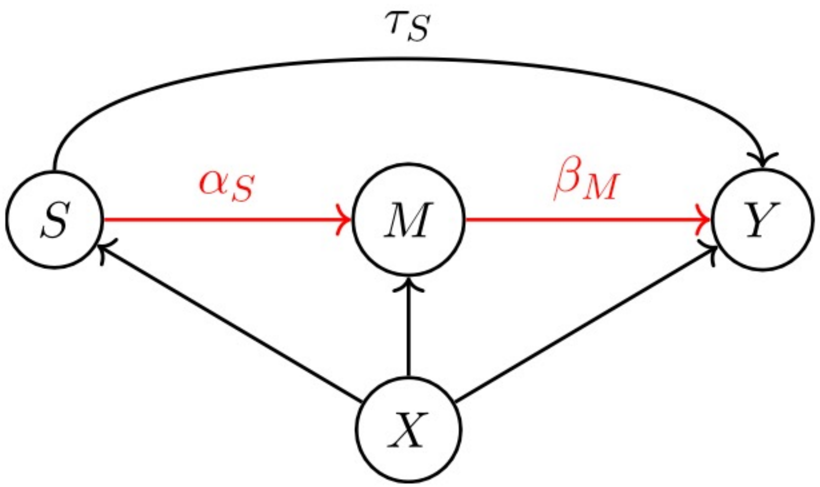
“Adaptive bootstrap tests for composite null hypotheses in the mediation pathway analysis”

Yinqiu He, Peter X.-K. Song and Gongjun Xu, 2023 JRSSB

Discussion by Zhiling Gu
Iowa State University, CIWG
November 28, 2023

Directed acyclic graph for mediation analysis



- *S*: Exposure
- *M*: Mediator
- *Y*: Outcome
- *X:* Potential confounders

"Total effect" of exposure S on Y
= **direct effect (not through M)** + **indirect effect (through M)**

Hypothesis test: Is there a mediator M, i.e. is there a mediator effect (ME)?

$$H_0 : \alpha_S \beta_M = 0 \text{ against } H_A : \alpha_s \neq 0 \text{ and } \beta_M \neq 0$$

Difficulty

$H_0 : \alpha_S\beta_M = 0$ is composed of three different parameter cases:
(i) $H_{0,1} : \alpha_S = 0$ and $\beta_M \neq 0$;
(ii) $H_{0,2} : \alpha_S \neq 0$ and $\beta_M = 0$; and
(iii) $H_{0,3} : \alpha_S = 0$ and $\beta_M = 0$.

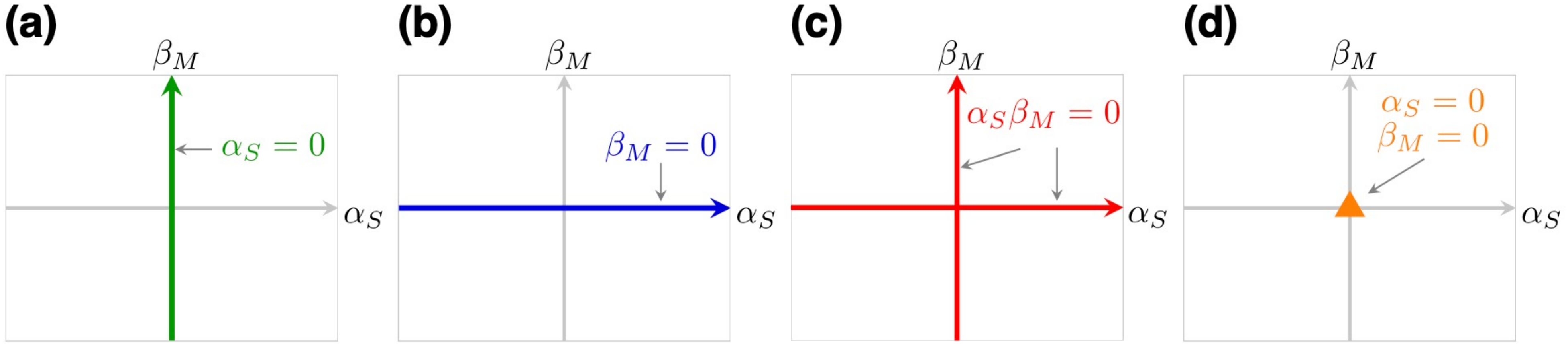


Figure 2. Visualisation of parameter spaces of (α_S, β_M) under different constraints. (a) $\alpha_S = 0$. (b) $\beta_M = 0$. (c) $\alpha_S\beta_M = 0$, and (d) $\alpha_S = \beta_M = 0$.

Note α_S and β_M are unknown and needs to be estimated, leading to **distinct asymptotic behaviours of test statistics**.

Existing Methods

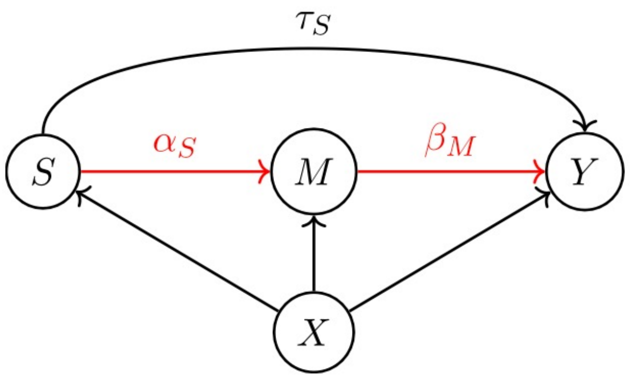
- Products of Coefficients (PoC)
  - **Sobel's test** (1982): Wald-type test and approximates the variance of $\hat{\alpha}_{S,n}, \hat{\beta}_{M,n}$
  - **Joint significance (JS) test** (Fritz & MacKinnon, 2007), also known as the MaxP test: rejects null hypothesis of no ME if both $\hat{\alpha}_{S,n}, \hat{\beta}_{M,n}$ pass a certain cut-off of statistical significance
  - ➢ **Overly conservative** in the neighborhood of $\left(\hat{\alpha}_S, \hat{\beta}_M\right) = (0,0)$

- Proposed Method: Adaptive bootstrap testing framework

Structural Equation Model (SEM)

Consider the popular linear SEM (MacKinnon, 2008; VanderWeele, 2015):

$$M = \alpha_S S + \boldsymbol{X}^\top \boldsymbol{a}_{\boldsymbol{X}} + \epsilon_M,$$
$$Y = \beta_M M + \boldsymbol{X}^\top \boldsymbol{\beta}_{\boldsymbol{X}} + \tau_S S + \epsilon_Y,$$



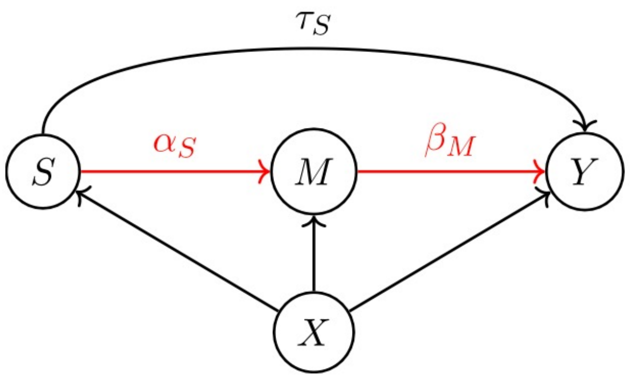
where

- $M(s)$ is the potential value of the mediator under exposure $S = s$
- $Y(s, m)$ is the potential outcome if exposure is set to $S = s$ and mediator is set to $M = m$
- *Stable unit treatment value (Rubin, 1980)*: $M = M(S), Y = Y(S, M(S))$
- n i.i.d observations, $\{(S_i, \mathbf{X}_i, M_i, Y_i), 1 \leq i \leq n\}$.

Structural Equation Model (SEM)

Consider the popular linear SEM (MacKinnon, 2008; VanderWeele, 2015):

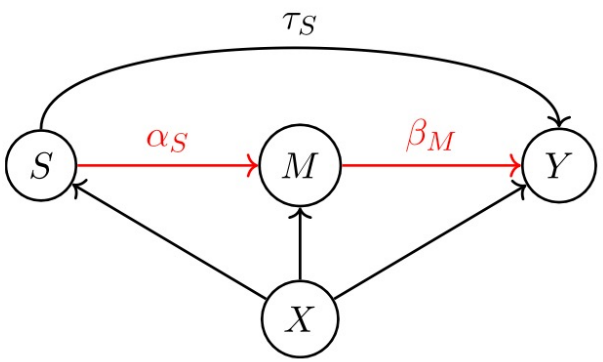
$$M = \alpha_S S + \boldsymbol{X}^\top \boldsymbol{a}_{\boldsymbol{X}} + \epsilon_M,$$
$$Y = \beta_M M + \boldsymbol{X}^\top \boldsymbol{\beta}_{\boldsymbol{X}} + \tau_S S + \epsilon_Y,$$



The ME or the natural indirect effect of $S = s$ vs. s^* (Imai et al., 2010) is defined as

$$\mathrm{E}\left\{Y(s, M(s)) - Y\left(s, M\left(s^*\right)\right)\right\}.$$

Assumptions



Assume that for all levels of s, s^*, and m

- $Y(s,m) \perp S \mid \{\boldsymbol{X} = \boldsymbol{x}\}$, no confounder for the relation of Y and S
- $Y(s,m) \perp M \mid \{S = s, \boldsymbol{X} = \boldsymbol{x}\}$, no confounder for the relation of Y and M conditioning on $S = s$
- $M(s) \perp S \mid \{\boldsymbol{X} = \boldsymbol{x}\}$, no confounder for the relation of M and S
- $Y(s,m) \perp M(s^*) \mid \{\boldsymbol{X} = \boldsymbol{x}\}$, no confounder for the $M - Y$ relation that is affected by S (VanderWeele & Vansteelandt, 2009)

Under these assumptions, the ME equals $\alpha_S \beta_M (s - s^*)$.

Non-regularity for simple SEM

Consider a simpler SEM:

$$M = \alpha_S S + \epsilon_M,$$
$$Y = \beta_M M + \epsilon_Y,$$

Classical asymptotic theory (van der Vaart, 2000):

$$\sqrt{n}\left(\hat{\alpha}_{S,n} - \alpha_S, \hat{\beta}_{M,n} - \beta_M\right)^\top \xrightarrow{d} \left(Z_{S,0}, Z_{M,0}\right)^\top$$

$$\sqrt{n}\left(\hat{\alpha}^*_{S,n} - \alpha_S, \hat{\beta}^*_{M,n} - \beta_M\right)^\top \xrightarrow{d} \left(Z'_{S,0}, Z'_{M,0}\right)^\top$$

where

- $\hat{\alpha}_{S,n}, \hat{\beta}_{M,n}$ are OLS estimators
- $\hat{\alpha}^*_{S,n}, \hat{\beta}^*_{M,n}$ are corresponding nonparametric bootstrap estimators

Non-regularity for simple SEM

Under $H_{0,3}$,

$$n(\hat{\alpha}_{S,n}\hat{\beta}_{M,n} - \alpha_S\beta_M) \xrightarrow{d} Z_{S,0}Z_{M,0},$$

and

$$\begin{aligned} & n(\hat{\alpha}^*_{S,n}\hat{\beta}^*_{M,n} - \hat{\alpha}_{S,n}\hat{\beta}_{M,n}) \\ =& n\{(\hat{\alpha}^*_{S,n} - \hat{\alpha}_{S,n})\hat{\beta}_{M,n} + (\hat{\beta}^*_{M,n} - \hat{\beta}_{M,n})\hat{\alpha}_{S,n} + (\hat{\alpha}^*_{S,n} - \hat{\alpha}_S)(\hat{\beta}^*_{M,n} - \hat{\beta}_{M,n})\} \\ \xrightarrow{d}& Z^{M,0}_{S,0} + Z_{S,0}Z'_{M,0} + Z^{M,0}_{S,0}, \end{aligned}$$

Inconsistency of classical nonparametric bootstrap

- Convergence rate n, not root n
- Limit of of bootstrap bias different from the limit of estimation bias

Non-regularity for local SEM

Consider local SEM:

$$M = \alpha_{S,n} S + \boldsymbol{X}^\top \boldsymbol{a}_{\boldsymbol{X}} + \epsilon_M,$$
$$Y = \beta_{M,n} M + \boldsymbol{X}^\top \boldsymbol{\beta}_{\boldsymbol{X}} + \tau_S S + \epsilon_Y,$$

where $\alpha_{S,n} = \alpha_S + n^{-1/2} b_\alpha$, and $\beta_{M,n} = \beta_M + n^{-1/2} b_\beta$ are locally perturbed counterparts of (α_S, β_M).

Condition 1 (C1.1) $\mathrm{E}(\epsilon_M | \boldsymbol{X}, S) = 0$ and $\mathrm{E}(\epsilon_Y | \boldsymbol{X}, S, M) = 0$. (C1.2) $\mathrm{E}(\boldsymbol{D}\boldsymbol{D}^\top)$ is a positive definite matrix with bounded eigenvalues, where $\boldsymbol{D} = (\boldsymbol{X}^\top, M, S)^\top$. (C1.3) The second moments of $(\epsilon_M, \epsilon_Y, S_\perp, M_{\perp'}, \epsilon_M S_\perp, \epsilon_Y M_{\perp'})$ are finite, where $S_\perp = S - \boldsymbol{X}^\top Q_{1,S}$ with $Q_{1,S} = \{\mathrm{E}(\boldsymbol{X}\boldsymbol{X}^\top)\}^{-1} \times \mathrm{E}(\boldsymbol{X}S)$, and $M_{\perp'} = M - \tilde{\boldsymbol{X}}^\top Q_{2,M}$ with $\tilde{\boldsymbol{X}} = (\boldsymbol{X}^\top, S)^\top$ and $Q_{2,M} = \{\mathrm{E}(\tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{X}})\}^{-1} \times \mathrm{E}(\tilde{\boldsymbol{X}} M)$.

Non-regularity for local SEM

Theorem 1 (Asymptotic Property). Assume Condition 1. Under the local model (6),

(i) when $(\alpha_S, \beta_M) \neq (0, 0)$, $\sqrt{n} \times (\hat{\alpha}_{S,n}\hat{\beta}_{M,n} - \alpha_{S,n}\beta_{M,n}) \xrightarrow{d} \alpha_S Z_M + \beta_M Z_S$;

(ii) when $(\alpha_S, \beta_M) = (0, 0)$, $n \times (\hat{\alpha}_{S,n}\hat{\beta}_{M,n} - \alpha_{S,n}\beta_{M,n}) \xrightarrow{d} b_\alpha Z_M + b_\beta Z_S + Z_M Z_S$,

where $(Z_S, Z_M)^\top$ is a mean-zero normal random vector with a covariance matrix given by that of the random vector $(\epsilon_M S_\perp / V_S, \epsilon_Y M_{\perp'} / V_M)^\top$ with $V_S = \mathrm{E}(S_\perp^2)$, and $V_M = \mathrm{E}(M_{\perp'}^2)$.

Proposal: Discern null cases because the asymptotic behavior of (i) and (ii) are different.

Idea: Isolate the possibility of (i) by comparing the absolute value of the standardized statistics $T_{\alpha,n} = \sqrt{n}\hat{\alpha}_{S,n}/\hat{\sigma}_{\alpha_S,n}$, $T_{\beta,n} = \sqrt{n}\hat{\beta}_{M,n}/\hat{\sigma}_{\beta_M,n}$ to some threshold.

Adaptive Bootstrap (AB) of PoC test

Strategy: Decompose the bias

$$\hat{\alpha}_{S,n}\hat{\beta}_{M,n} - \alpha_{S,n}\beta_{M,n}$$
$$= \underbrace{(\hat{\alpha}_{S,n}\hat{\beta}_{M,n} - \alpha_{S,n}\beta_{M,n}) \times (1 - I_{\alpha_S,\lambda_n} I_{\beta_M,\lambda_n})}_{(i)} + \underbrace{(\hat{\alpha}_{S,n}\hat{\beta}_{M,n} - \alpha_{S,n}\beta_{M,n}) \times I_{\alpha_S,\lambda_n} I_{\beta_M,\lambda_n}}_{(ii)}$$

where $I_{\alpha_S,\lambda_n} = I\{|T_{\alpha,n}| \leq \lambda_n, \alpha_S = 0\}$ and $I_{\beta_M,\lambda_n} = I\{|T_\beta, n| \leq \lambda_n, \beta_M = 0\}$

Notations:

- P_n: population probability measure of $(S, \boldsymbol{X}, M, Y)$
- $\mathbb{P}_n$: empirical probability measure of $\{(S_i, \boldsymbol{X}_i, M_i, Y_i), 1 \leq i \leq n\}$
- $\mathbb{P}_n^*$: non-parametric bootstrap version of $\mathbb{P}_n$
- $f(S, \boldsymbol{X}, M, Y)$: any measurable function
- $\mathbb{G}_n f = \sqrt{n}(\mathbb{P}_n - P_n) f = \sqrt{n}\{n^{-1}\sum_{i=1}^n f(S_i, \boldsymbol{X}_i, M_i, Y_i) - \mathrm{E} f(S, \boldsymbol{X}, M, Y)\}$: empirical process; $\mathbb{G}_n^*$: non-parametric bootstrap counterpart of $\mathbb{G}_n$

Adaptive Bootstrap (AB) of PoC test

Proposed Statistic:

$$U^* = \underbrace{(\hat{\alpha}^*_{S,n}\hat{\beta}^*_{M,n} - \hat{\alpha}_{S,n}\hat{\beta}_{M,n}) \times (1 - I^*_{\alpha_S,\lambda_n} I^*_{\beta_M,\lambda_n})}_{(i)} + \underbrace{n^{-1}\mathbb{R}^*_n(b_\alpha, b_\beta) \times I^*_{\alpha_S,\lambda_n} I^*_{\beta_M,\lambda_n}}_{(ii)}$$

where $I^*_{\alpha_S,\lambda_n} = I\{|T^*_{\alpha,n}| \le \lambda_n, |T_{\alpha,n}| \le \lambda_n\}$, $I^*_{\beta_M,\lambda_n} = I\{|T^*_{\beta,n}| \le \lambda_n, |T_{\beta,n}| \le \lambda_n\}$,

- $\mathbb{R}^*_n(b_\alpha, b_\beta) := b_\alpha \mathbb{Z}^*_{M,n} + b_\beta \mathbb{Z}^*_{S,n} + \mathbb{Z}^*_{M,n}\mathbb{Z}^*_{S,n}$
- $\mathbb{Z}^*_{S,n} = \mathbb{G}^*_n(\hat{\varepsilon}_{M,n} S^*_\perp)/V^*_{S,n}$
- $\mathbb{Z}^*_{M,n} = \mathbb{G}^*_n(\hat{\varepsilon}_{Y,n} M^*_{\perp'})/V^*_{M,n}$
- $\hat{S}_\perp = S - \boldsymbol{X}^\top\{\mathbb{P}_n(\boldsymbol{X}\boldsymbol{X}^\top)\}^{-1}\mathbb{P}_n(\boldsymbol{X}S)$
- $\hat{M}_{\perp'} = M - \tilde{\boldsymbol{X}}^\top\{\mathbb{P}_n(\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^\top)\}^{-1}\mathbb{P}_n(\tilde{\boldsymbol{X}}M)$, $\tilde{\boldsymbol{X}} = (\boldsymbol{X}^\top, S)^\top$
- $\mathbb{V}^*_{S,n} = \mathbb{P}^*_n\{(S^*_\perp)^2\}$, $\mathbb{V}^*_{M,n} = \mathbb{P}^*_n\{(M^*_{\perp'})^2\}$

$T_{\alpha,n} = \sqrt{n}\hat{\alpha}_{S,n}/\hat{\sigma}_{\alpha_{S,n}}$
$T_{\beta,n} = \sqrt{n}\hat{\beta}_{M,n}/\hat{\sigma}_{\beta_{M,n}}$

Adaptive Bootstrap (AB) of PoC test

Theorem 2 (Adaptive Bootstrap Consistency). Assume the conditions of Theorem 1 are satisfied. When $\lambda_n = o(\sqrt{n})$ and $\lambda_n \to \infty$ as $n \to \infty$,

$$c_n U^* \overset{d^*}{\rightsquigarrow} c_n(\hat{\alpha}_{S,n}\hat{\beta}_{M,n} - \alpha_S\beta_M),$$

where c_n is a non-random scaling factor satisfying

$$c_n = \begin{cases} \sqrt{n}, & \text{when } (\alpha_S, \beta_M) \neq (0,0) \\ n, & \text{when } (\alpha_S, \beta_M) = (0,0). \end{cases} \tag{9}$$

Consistency of proposed nonparametric adaptive bootstrap

- Proper scaling regardless of the underlying true null

Test Procedure:

- Given a nominal level ω let $q(\omega/2)$ and $q(1-\omega/2)$ denote the lower and upper $\omega/2$ quantiles, respectively, of the bootstrap estimates U^*.
- If $\hat{\alpha}_{S,n}\hat{\beta}_{M,n}$ falls outside the interval $(q(\omega/2), q(1-\omega/2))$, we reject the composite null, and conclude that the ME is statistically significant at the level ω.

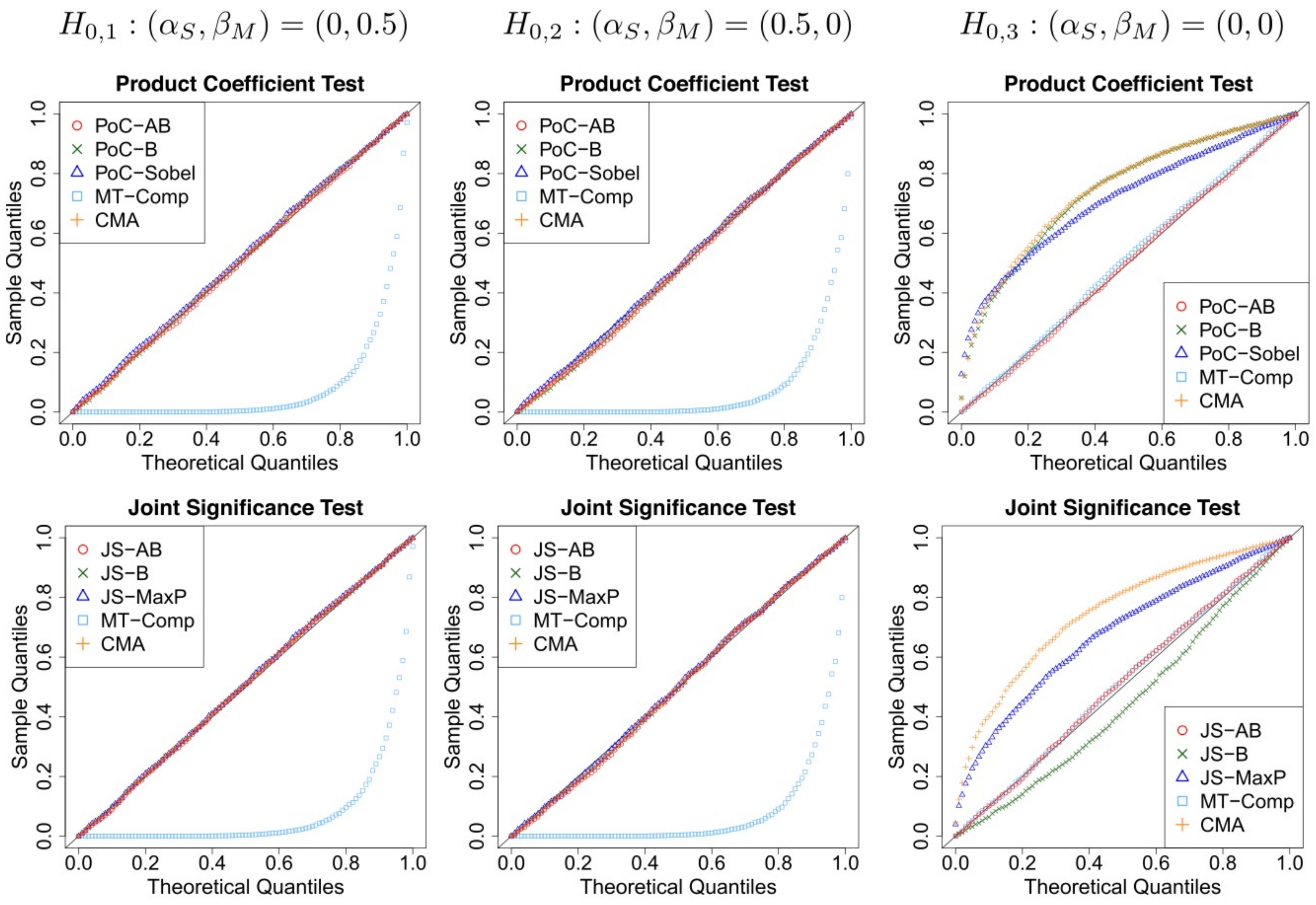


Figure 4. Q–Q plots of p-values under the fixed null with $n = 200$.

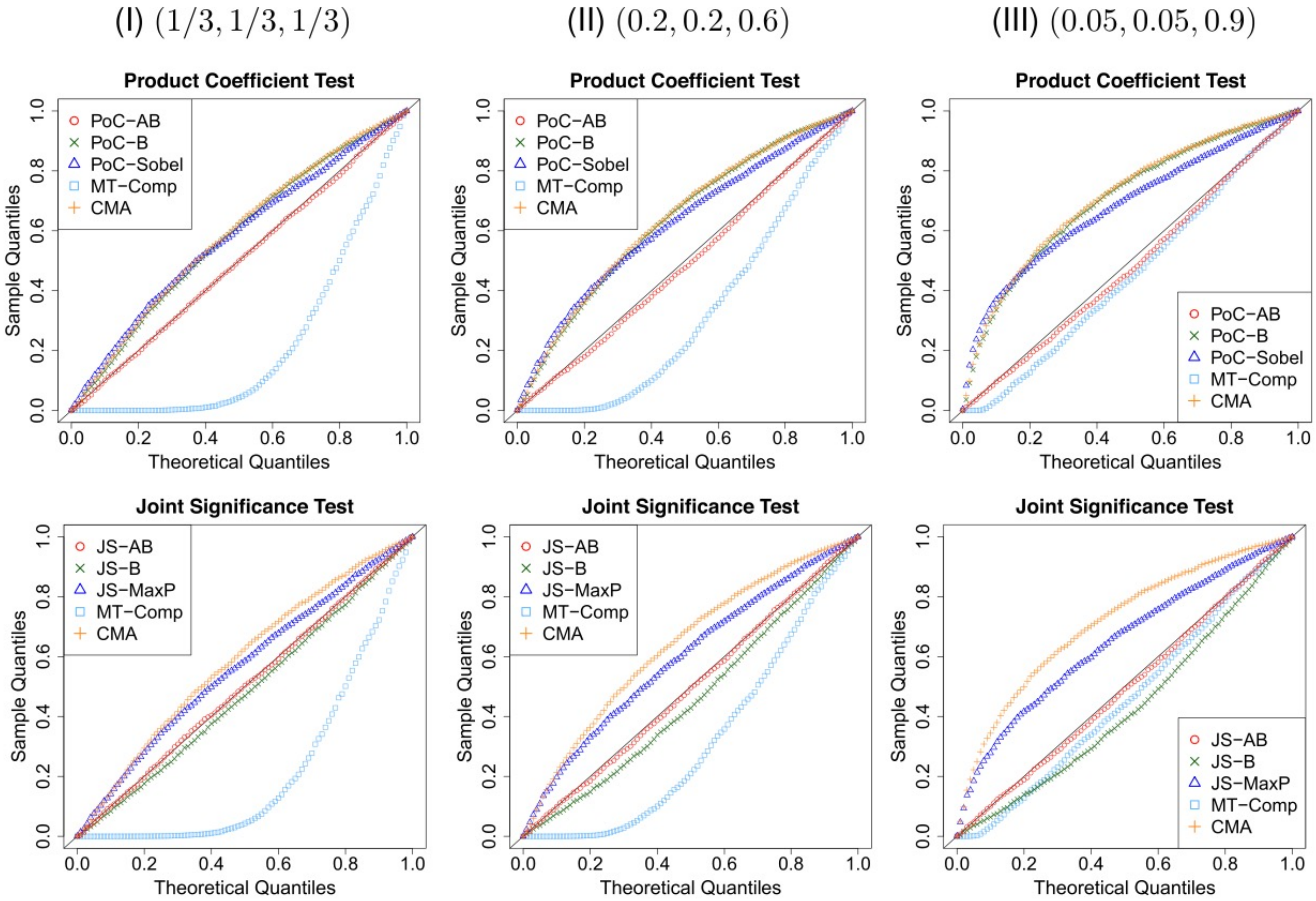


Figure 5. Q–Q plots of p-values under the mixture of nulls: $n = 200$.

Thank you!