Supplementary Material for "Total-effect Test can Erroneously Reject Complete Mediation — Proofs and Examples"

Tingxuan Han[1], Luxi Zhang[2], Dianshi Li[2], Xinshu Zhao[2*], Ke Deng[1*]

[1*]Department of Statistics and Data Science, Tsinghua University, Beijing, China.
[2]Department of Communication, University of Macau, Macau, China.

*Corresponding author(s). E-mail(s): xszhao@um.edu.mo; kdeng@tsinghua.edu.cn;
Contributing authors: htx21@mails.tsinghua.edu.cn; yc27303@um.edu.mo; yc37228@um.edu.mo;

Section S1 presents the detailed proof for Lemma 1, 2, 4, Theorem 3, 6, 7 and 8. Section S2 shows the results of simulation for competitive mediation under LSE-F, LSE-Sobel and LAD-Z frameworks. Section S3 gives a brief review of the history of the percentage contribution to total effect t_p and adds our analysis.

S1 Technical Proofs

S1.1 Proof of Lemma 1

The invariance of F-test has been proved in Jiang et al. (2021) and we focus on the invariance of Sobel test here. To simplify the problem, consider the following general multivariate linear regression problem:

$$Y = \beta_0 X_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon$$

with n observed data points $\{(X_{i0}, X_{i1}, \ldots, X_{ip}, Y_i)\}$. Define the vector of coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)'$, design matrix $\boldsymbol{X} = (\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_p)$, where $\boldsymbol{X}_j = (X_{1j}, \ldots, X_{nj})'$, and response vector $\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$. Denote the 2-norm of a vector as $||\cdot||_2^2$. Then the least square estimator of $\boldsymbol{\beta}$ takes the form of

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y},$$

and the covariance is

$$\text{cov}(\hat{\boldsymbol{\beta}}) = s^2(\boldsymbol{X}'\boldsymbol{X})^{-1},$$

where $s^2 = ||\boldsymbol{Y} - \boldsymbol{Y}_{\boldsymbol{X}}||_2^2/(n-p-1)$, $\boldsymbol{Y}_{\boldsymbol{X}}$ represents projection of $\boldsymbol{Y}$ onto the space spanned by $\boldsymbol{X}$. As Sobel test is based on $\hat{\boldsymbol{\beta}}$ and $\text{cov}(\hat{\boldsymbol{\beta}})$, it suffices to show that for orthogonal matrix Γ and constant γ, the LSE and estimated covariance of coefficient of the regression problem with transformed data matrix $(\tilde{\boldsymbol{X}}_0, \ldots, \tilde{\boldsymbol{X}}_p, \tilde{\boldsymbol{Y}}) = \gamma\Gamma(\boldsymbol{X}_0, \ldots, \boldsymbol{X}_p, \boldsymbol{Y})$ remain unchanged.

First, the invariance of LSE is easy to see as

$$\tilde{\boldsymbol{\beta}} = (\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{Y}} = (\boldsymbol{X}'\Gamma'\Gamma\boldsymbol{X})^{-1}\boldsymbol{X}'\Gamma'\Gamma\boldsymbol{Y} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} = \hat{\boldsymbol{\beta}}.$$

Moreover, since

$$||\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{Y}}_{\tilde{\boldsymbol{X}}}||_2^2 = ||\gamma\Gamma(\boldsymbol{Y} - \boldsymbol{Y}_{\boldsymbol{X}})||_2^2 = \gamma^2||\boldsymbol{Y} - \boldsymbol{Y}_{\boldsymbol{X}}||_2^2,$$

and

$$(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1} = (\gamma^2\boldsymbol{X}'\Gamma'\Gamma\boldsymbol{X})^{-1} = (\gamma^2\boldsymbol{X}'\boldsymbol{X})^{-1},$$

we have

$$\text{cov}(\tilde{\boldsymbol{\beta}}) = ||\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{Y}}_{\tilde{\boldsymbol{X}}}||_2^2/(n-2)\cdot(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1} = s^2(\boldsymbol{X}'\boldsymbol{X})^{-1} = \text{cov}(\hat{\boldsymbol{\beta}}).$$

Above all, the LSE and estimated covariance are invariant under orthogonal transformation. Hence, the invariance of LSE-Sobel test holds as well.

S1.2 Proof of Lemma 2

By definition, the test statistics in Sobel test is

$$S = \frac{\hat{a}\hat{b}}{\left(\hat{a}^2\text{Var}(\hat{b}) + \hat{b}^2\text{Var}(\hat{a})\right)^{1/2}} = \frac{T_aT_b}{\sqrt{T_a^2 + T_b^2}},$$

where $T_a^2 = \hat{a}^2/\text{Var}(\hat{a})$ and $T_b^2 = \hat{b}^2/\text{Var}(\hat{b})$. Based on the transformed data matrix $\tilde{\mathcal{D}}$, we have

$$\hat{a} = \frac{m_2}{x_2}, \quad \hat{b} = \frac{y_3}{m_3}, \quad \text{Var}(\hat{a}) = \frac{m_3^2}{(n-2)x_2^2}, \quad \text{Var}(\hat{b}) = \frac{y_4^2}{(n-3)m_3^2}.$$

Hence, $T_a = \sqrt{n-2}\cdot m_2/m_3$, $T_b = \sqrt{n-3}\cdot y_3/y_4$, and

$$|S| = \frac{1}{(1/T_a^2 + 1/T_b^2)^{1/2}} = \frac{1}{\{1/[(n-2)r^2] + 1/[(n-3)p^2]\}^{1/2}}.$$

As sample size $n \to \infty$, S follows standard normal distribution asymptotically. Hence, the rejection region of $a \times b$ is

$$\mathcal{R}_{a\times b}(\alpha) = \left\{|S| > z_{\alpha/2}\right\} = \left\{\frac{1}{(n-2)r^2} + \frac{1}{(n-3)p^2} < \frac{1}{z_{\alpha/2}^2}\right\}.$$

S1.3 Proof of Theorem 3

As the geometric plots of rejection regions have been shown in Jiang et al. (2021), we only give the mathematical analysis here.

Let $\mathcal{R}_\beta(\alpha \mid r)$ be the intersection of $\mathcal{R}_\beta(\alpha)$ and the p-q plane $\mathcal{P}_r$ for all $\beta \in \{a, b, d, c, a\times b\}$. It is straightforward to see that $\mathcal{R}_\beta(\alpha) = \bigcup_r \mathcal{R}_\beta(\alpha \mid r)$, and

$$\bar{\mathcal{R}}_c(\alpha) \cap \mathcal{R}_{a\times b}(\alpha) \cap \mathcal{R}_d(\alpha) = \bigcup_{r>0} \left\{\bar{\mathcal{R}}_c(\alpha \mid r) \cap \mathcal{R}_{a\times b}(\alpha \mid r) \cap \mathcal{R}_d(\alpha \mid r)\right\}.$$

As implied by Lemma 2, $\mathcal{R}_{a\times b}(\alpha|r) = \{p > p_0(r)\}$, where

$$\frac{1}{(n-2)r^2} + \frac{1}{(n-3)p_0^2(r)} = \frac{1}{z_{\alpha/2}^2}, \ (n-2)r^2 > z_{\alpha/2}^2. \tag{S1}$$

By Jiang et al. (2021), $\hat{a}\hat{b}\hat{d} > 0$ implies $\hat{a}\hat{b}\hat{c} > 0$ and $\mathcal{R}_d(\alpha \mid r) = \{q > rp + p_{n,\alpha}(1 + r^2)^{1/2}\}$. Then for $r^2 > z_{\alpha/2}^2/(n-2)$, the intersection of $\mathcal{R}_{a\times b}(\alpha \mid r)$, $\mathcal{R}_d(\alpha \mid r)$ and $\bar{\mathcal{R}}_c(\alpha \mid r)$ is

$$\left\{p > p_0(r), rp + p_{n,\alpha}(1+r^2)^{1/2} < q \le r_{n,\alpha}(p^2+1)^{1/2}\right\}.$$

Hence, it suffices to show that as $n \to \infty$,

$$P\left(\mathcal{D} \in \bigcup_{r>r_0(n)} \left\{p > p_0(r), rp + p_{n,\alpha}(1+r^2)^{1/2} < q \le r_{n,\alpha}(p^2+1)^{1/2}\right\}\right) \to 0, \tag{S2}$$

where $r_0(n) = z_{\alpha/2}/\sqrt{n-2}$.

Since the probability density function of F-distribution with degrees of freedom $(1, n)$ satisfies

$$f_{F_{1,n}}(x) = \frac{\Gamma(n/2 + 1/2)}{\sqrt{\pi}\Gamma(n/2)} \cdot \frac{1}{\sqrt{nx}} \left(1 + \frac{x}{n}\right)^{-\frac{n+1}{2}} \to \frac{1}{\sqrt{2\pi x}} e^{-x/2}$$

as $n \to \infty$, we have $\lambda_{n-2}(\alpha) \to \chi_1^2(\alpha)$, where $\chi_1^2(\alpha)$ is the αth-quantile of χ_1^2 distribution. Hence, $\sqrt{\lambda_{n-2}(\alpha)} \to \sqrt{\chi_1^2(\alpha)} \equiv z_{\alpha/2}$, i.e., $\sqrt{n-2} \cdot r_{n,\alpha} \equiv \sqrt{\lambda_{n-2}(\alpha)} \to z_{\alpha/2} = \sqrt{n-2} \cdot r_0(n)$. Therefore, $\{r : r_0(n) < r < r_{n,\alpha}\} \to \emptyset$. According to Jiang et al. (2021),

$$\bigcup_{r \geq r_{n,\alpha}} \left\{p > 0 : rp + p_{n,\alpha}(1+r^2)^{1/2} < r_{n,\alpha}(p^2+1)^{1/2}\right\} = \emptyset.$$

Hence,

$$\begin{aligned} & P\left(\mathcal{D} \in \bigcup_{r > r_0(n)} \left\{p > p_0(r), rp + p_{n,\alpha}(1+r^2)^{1/2} < r_{n,\alpha}(p^2+1)^{1/2}\right\}\right) \\ & = P\left(\mathcal{D} \in \bigcup_{r_0(n) < r < r_{n,\alpha}} \left\{p > p_0(r), rp + p_{n,\alpha}(1+r^2)^{1/2} < r_{n,\alpha}(p^2+1)^{1/2}\right\}\right) \\ & \to 0, \end{aligned}$$

and the argument (S2) holds.

S1.4 Proof of Lemma 4

By Lemma 2,

$$\hat{a}\hat{b}\hat{d} = \frac{m_2 y_3(m_3 y_2 - m_2 y_3)}{x_2^2 m_3^2}, \quad \hat{a}\hat{b}\hat{c} = \frac{m_2 y_2 y_3}{x_2^2 m_3}.$$

Since $x_2 > 0$ and $m_3 > 0$, conditions $\hat{a}\hat{b}\hat{d} < 0$ and $\hat{a}\hat{b}\hat{c} \geq 0$ imply that

$$m_2 m_3 y_2 y_3 < m_2^2 y_3^2 \text{ and } m_2 y_2 y_3 \geq 0.$$

Hence, $m_2 m_3 y_2 y_3 > 0$ and $|m_3 y_2| < |m_2 y_3|$, which is equivalent to $q < rp$ as $y_4 > 0$.

S1.5 Proof of Theorem 6

We show that there exists $N > 0$ such that for any $n > N$, we can find some $r_n > 0$ such that $\bar{\mathcal{R}}_c(\alpha|r_n) \cap \mathcal{R}_{a\times b}(\alpha|r_n) \cap \bar{\mathcal{R}}_d(\alpha|r_n) \neq \emptyset$. When $\hat{a}\hat{b}\hat{c} \geq 0$, $\mathcal{R}_{a\times b}(\alpha \mid r) \cap \bar{\mathcal{R}}_c(\alpha \mid r) \cap \bar{\mathcal{R}}_d(\alpha \mid r)$ takes the form

$$\left\{\max\left\{rp - p_{n,\alpha}(r^2+1)^{1/2}, 0\right\} \leq q \leq \min\{r_{n,\alpha}(p^2+1)^{1/2}, rp + p_{n,\alpha}(r^2+1)^{1/2}\}, p > p_0(r)\right\},$$

and when $\hat{a}\hat{b}\hat{c} < 0$, the above intersection is

$$\left\{0 \leq q \leq \min\left\{r_{n,\alpha}(p^2+1)^{1/2}, -rp + p_{n,\alpha}(r^2+1)^{1/2}\right\}, p > p_0(r)\right\}.$$

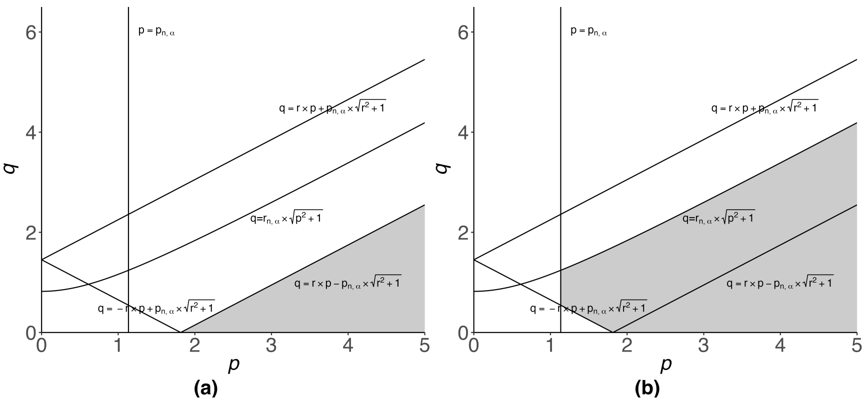

Fig. S1 Geometry of $\mathcal{R}_a(\alpha \mid r) \cap \mathcal{R}_b(\alpha \mid r) \cap \mathcal{R}_d(\alpha \mid r) \cap \mathcal{R}_c^c(\alpha \mid r)$ in p-q plane with $r > r_{n,\alpha}$ for competitive mediation when (a) $\hat{a}\hat{b}\hat{c} \geq 0$ and (b) $\hat{a}\hat{b}\hat{c} < 0$.

It is easy to see that $\bar{\mathcal{R}}_c(\alpha|r) \cap \mathcal{R}_{a\times b}(\alpha|r) \cap \bar{\mathcal{R}}_d(\alpha|r) \neq \emptyset$ holds if $p_0(r) < p_{n,\alpha}(r^2 + 1)^{1/2}/r$. As the definition of $p_0(r)$ implies $r > z_{\alpha/2}/\sqrt{n-2}$, it suffices to show that there exists $N > 0$ such that for any $n > N$, we can find some $r_n > z_{\alpha/2}/\sqrt{n-2}$ such that

$$p_0(r_n) < p_{n,\alpha}(r_n^2 + 1)^{1/2}/r_n. \tag{S3}$$

Let $r_n^2 = 2z_{\alpha/2}^2/(n-2)$, then Eq. (S1) implies $p_0^2(r_n) = 2z_{\alpha/2}^2/(n-3)$. Since $\sqrt{n-3} \cdot p_{n,\alpha} \to z_{\alpha/2}$ and $r_n^2 \to 0$, there exists $N > 0$ s.t. for $n > N$, we have $\sqrt{n-3} \cdot p_{n,\alpha} \geq z_{\alpha/2}/2$ and $r_n < 1/3$. Therefore, when $n > N$, we have

$$\sqrt{n-3} \cdot p_{n,\alpha}(r_n^2 + 1)^{1/2}/r_n \geq \sqrt{10} \cdot z_{\alpha/2}/2 > \sqrt{2} z_{\alpha/2} = \sqrt{n-3} \cdot p_0(r_n).$$

Above all, for $n > N$ and $r_n^2 = 2z_{\alpha/2}^2/(n-2)$, we have $\bar{\mathcal{R}}_c(\alpha|r_n) \cap \mathcal{R}_{a\times b}(\alpha|r_n) \cap \bar{\mathcal{R}}_d(\alpha|r_n) \neq \emptyset$, and thus, $\bar{\mathcal{R}}_c(\alpha) \cap \mathcal{R}_{a\times b}(\alpha) \cap \bar{\mathcal{R}}_d(\alpha) \neq \emptyset$.

S1.6 Proof of Theorem 7

$\mathcal{R}_a(\alpha) \cap \mathcal{R}_b(\alpha) \cap \mathcal{R}_d(\alpha) \cap \bar{\mathcal{R}}_c(\alpha) \neq \emptyset$ can be equivalently expressed as $\mathcal{R}_a(\alpha \mid r) \cap \mathcal{R}_b(\alpha \mid r) \cap \mathcal{R}_d(\alpha \mid r) \cap \bar{\mathcal{R}}_c(\alpha \mid r) \neq \emptyset$ for some $r > r_{n,\alpha}$ since $\mathcal{R}_a(\alpha \mid r) = \mathcal{P}_r \cap \{r > r_{n,\alpha}\} = \emptyset$ when $r \leq r_{n,\alpha}$. By definition, $\mathcal{R}_b(\alpha \mid r) = \{p > p_{n,\alpha}\}$, and $\bar{\mathcal{R}}_c(\alpha \mid r) = \left\{p \geq 0, 0 \leq q \leq r_{n,\alpha}(p^2+1)^{1/2}\right\}$. When $\hat{a}\hat{b}\hat{c} < 0$,

$$\mathcal{R}_d(\alpha \mid r) = \left\{p \geq 0, q > -rp + p_{n,\alpha}(r^2+1)^{1/2}\right\},$$

and for $\hat{a}\hat{b}\hat{c} \geq 0$, Lemma 4 implies that

$$\mathcal{R}_d(\alpha \mid r) = \left\{p \geq 0, 0 \leq q < rp - p_{n,\alpha}(r^2+1)^{1/2}\right\}.$$

The geometry of $\mathcal{R}_a(\alpha \mid r) \cap \mathcal{R}_b(\alpha \mid r) \cap \mathcal{R}_d(\alpha \mid r) \cap \bar{\mathcal{R}}_c(\alpha \mid r)$ is demonstrated in Figure S1. When $\hat{a}\hat{b}\hat{c} \geq 0$, it suffices to show that $r_{n,\alpha}(p^2+1)^{1/2} > \max\{0, rp - p_{n,\alpha}(r^2+1)^{1/2}\}$ for some $p > p_{n,\alpha}$, which always holds for $p \in (p_{n,\alpha}, p_{n,\alpha}(r^2+1)^{1/2}/r)$. When $\hat{a}\hat{b}\hat{c} < 0$, it suffices to show that $r_{n,\alpha}(p^2+1)^{1/2} > 0$ for some $p > p_{n,\alpha}$, which always holds. We thereby conclude the proof of Theorem 7.

As shown in Figure S1, the intersection of $\mathcal{R}_{a\times b}(\alpha)$, $\mathcal{R}_d(\alpha)$ and $\bar{\mathcal{R}}_c(\alpha)$ under $\hat{a} \times \hat{b} \times \hat{c} \geq 0$ is a subset of that under $\hat{a} \times \hat{b} \times \hat{c} < 0$, which implies the total-effect test is more likely to be erroneous for establishing competitive mediation when $\hat{a} \times \hat{b} \times \hat{c} < 0$.

S1.7 Proof of Theorem 8

We show that $\bar{\mathcal{R}}_c(\alpha|r) \cap \mathcal{R}_{a\times b}(\alpha|r) \cap \mathcal{R}_d(\alpha|r) \neq \emptyset$ for some r when $\hat{a}\hat{b}\hat{d} < 0$. We only consider the case where $r > z_{\alpha/2}/\sqrt{n-2}$ since $\mathcal{R}_{a\times b}(\alpha|r) = \emptyset$ when this condition doesn't hold. The expression of $\mathcal{R}_d(\alpha|r)$ is the same as that under the LSE-F framework. As $\bar{\mathcal{R}}_c(\alpha|r) = \{p \geq 0, q \leq r_{n,\alpha}(p^2+1)^{1/2}\}$ and $\mathcal{R}_{a\times b}(\alpha|r) = \{p > p_0(r)\}$, the observation that $\bar{\mathcal{R}}_c(\alpha|r) \cap \mathcal{R}_{a\times b}(\alpha|r) \cap \mathcal{R}_d(\alpha|r) \neq \emptyset$ is trivial.

S2 Simulations for Competitive Mediation

This section presents the results of Monte Carlo simulation for competitive mediation under LSE-F, LSE-Sobel and LAD-Z frameworks.

To validate Theorem 7, we generate the simulated data from model (1) and (2) as follows:

$$n \sim \text{Unif}(\{10, \ldots, 100\}), \quad (i_M, i_Y, a, b, d) \sim \text{Unif}[-1,1]^5,$$

$$X \sim N(0,1), \quad \sigma_M^2 \text{ and } \sigma_Y^2 \sim \text{Inv-Gamma}(1,1).$$

A total of $10,000$ independent datasets of different sample sizes were simulated. For each simulated dataset, the LSEs $(\hat{a}, \hat{b}, \hat{c}, \hat{d})$ and their p-values (p_a, p_b, p_c, p_d) under the LSE-F framework are calculated. If Theorem 7 holds, then when for any fixed $\alpha \in (0,1)$, we have $\{p_c \geq \alpha\} \cap \{\hat{a}\hat{b}\hat{d} < 0\} \neq \emptyset$ if $\max(p_a, p_b, p_d) < \alpha$.

Figure S2 (a) checks the p-value condition when $\alpha = 0.1$ by demonstrating each simulated dataset with one point in a 2-dimensional space with $\max(p_a, p_b, p_d)$ and p_c be the X and Y-axis, respectively. The solid circles stand for datasets satisfying $\max(p_a, p_b, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$, gray crossings represent data sets such that $\max(p_a, p_b, p_d) \geq \alpha$ or $\hat{a}\hat{b}\hat{d} \geq 0$, and the dark gray dashed line represents $p_c = \alpha$. Solid circles above the line $p_c = \alpha$ form an empirical version of set $\{p_c \geq \alpha\}$. We can see that when $\max(p_a, p_b, p_d) < \alpha = 0.1$, $\{p_c \geq \alpha\} \cap \{\hat{a}\hat{b}\hat{d} < 0\}$ is not empty. To check the theoretical result for different values of α, the proportion of datasets satisfying $p_c \geq \alpha$ for 1000 evenly spaced values of α in $(0.01, 0.99)$ when $\max(p_a, p_b, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$ under LSE-F framework is shown in Figure S2 (b). It implies that the total-effect test will reject competitive mediation erroneously with quite large probability.

Similar analysis under LSE-Sobel framework and LAD-Z framework could be conducted to test whether a similar result holds for other frameworks for establishing competitive mediation. For LSE-Sobel framework, we calculated the LSEs $(\hat{a}, \hat{b}, \hat{c}, \hat{d})$

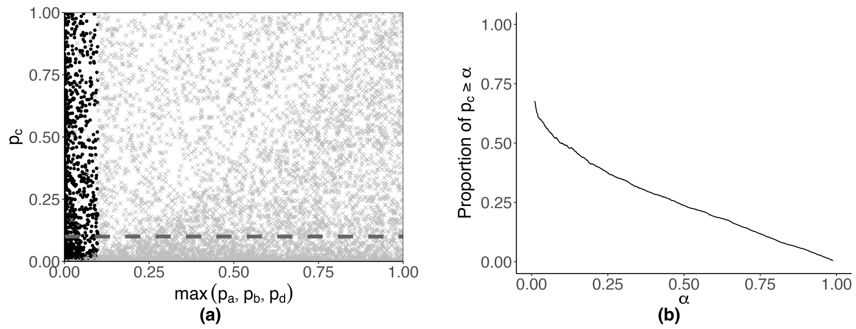

Fig. S2 (a) Scatter plot of p-values with $\alpha = 0.1$ under LSE-F framework: black solid circles represent datasets with $\max(p_a, p_b, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$, grey crossings represent datasets with $\max(p_a, p_b, p_d) \geq \alpha$ or $\hat{a}\hat{b}\hat{d} \geq 0$, and the dark gray dashed line represents $p_c = \alpha$. (b) Proportion of datasets satisfying $p_c \geq \alpha$ for different α when $\max(p_a, p_b, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$ under LSE-F framework.

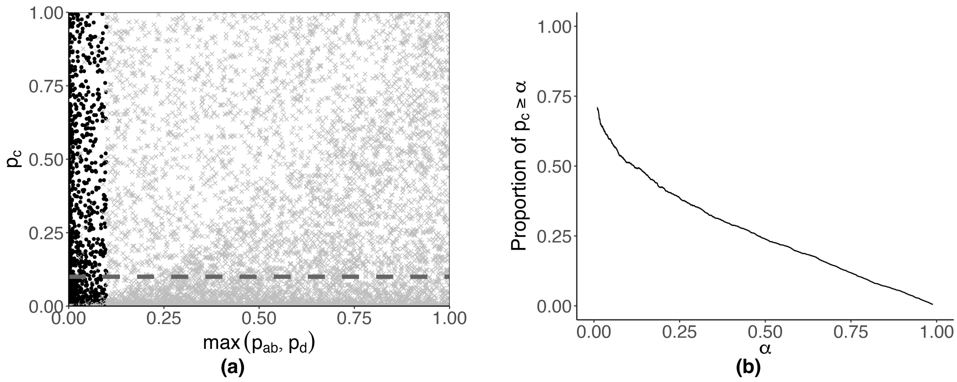

Fig. S3 (a) Scatter plot of p-values with $\alpha = 0.1$ under LSE-Sobel framework: black solid circles represent datasets with $\max(p_{ab}, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$, grey crossings represent datasets with $\max(p_{ab}, p_d) \geq \alpha$ or $\hat{a}\hat{b}\hat{d} \geq 0$, and the dark gray dashed line represents $p_c = \alpha$. (b) Proportion of datasets satisfying $p_c \geq \alpha$ for different α when $\max(p_{ab}, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$ under LSE-Sobel framework.

for each simulated dataset using the same group of simulated datasets. p-values (p_c, p_d) of F-test as well as the p-value p_{ab} of Sobel test for $a \times b$ are calculated. If a similar result holds for LSE-Sobel framework, we could expect to see for any fixed $\alpha \in (0, 1)$, when $\max(p_{ab}, p_d) < \alpha$, we have $\{p_c \geq \alpha\} \cap \{\hat{a}\hat{b}\hat{d} < 0\} \neq \emptyset$, which is supported by the results in Figure S3.

For LAD-Z framework, The LAD estimator $\hat{a}, \hat{b}, \hat{d}, \hat{c}$ as well as their corresponding p-values under Z-test are calculated. Similarly, If $\{p_c \geq \alpha\} \cap \{\hat{a}\hat{b}\hat{d} < 0\} \neq \emptyset$ for any fixed $\alpha \in (0, 1)$ when $\max(p_a, p_b, p_d) < \alpha$, the same conclusion can be reached under LAD-Z framework. Results are shown in Figure S4.

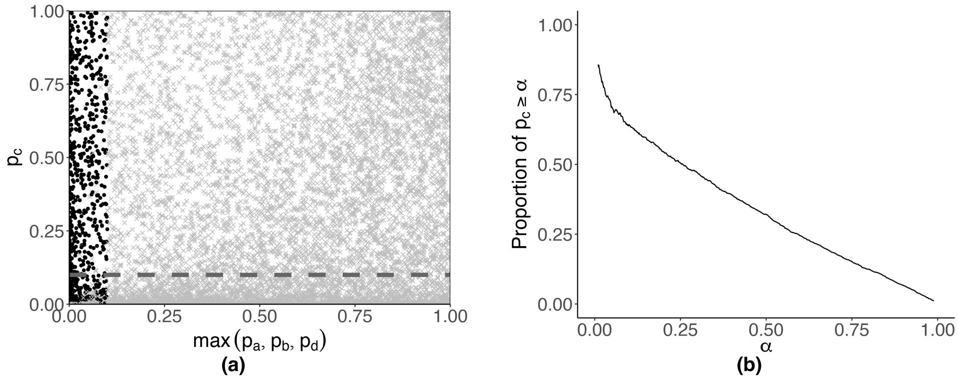

Fig. S4 (a) Scatter plot of p-values with $\alpha = 0.1$ under LAD-Z framework: black solid circles represent datasets with $\max(p_a, p_b, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$, grey crossings represent datasets with $\max(p_a, p_b, p_d) \geq \alpha$ or $\hat{a}\hat{b}\hat{d} \geq 0$, and the dark gray dashed line represents $p_c = \alpha$. (b) Proportion of datasets satisfying $p_c \geq \alpha$ for different α when $\max(p_a, p_b, p_d) < \alpha$ and $\hat{a}\hat{b}\hat{d} < 0$ under LAD-Z framework.

S3 Percentage Contribution to Total Effect

To help dissect the product, discern the parts, and divine the process, PAPA calculates *percentage contribution to total effect* (t_p), the contribution of each elemental part, i.e., the a, b, ab, d, or c path, to the $IV \to DV$ total effect, c, as detailed in Table 2. This section gives a brief review of the history of the ratio and adds our analysis.

S3.1 Proportion Mediated vs Percentage Contribution

S3.1.1 Proportion Mediated

Alwin and Hauser (1975) introduced the core concept of the ratio, which they named the proportion of *the total effect mediated by* X and *the proportion of the total effect unmediated by* X (p. 43). Sobel (1982) recommended it under the same label (p. 305). MacKinnon and Dwyer (1993) recommended it after renaming it *the percentage of the total effect that is mediated* (p.152). Ditlevsen et al. (2005) advocated it as *mediation proportion*. Huang et al. (2004) applied it to logistic mediation model under the label *relative indirect effect* (RIND). Freedman et al. (1992) and Freedman (2001) raised its significance by renaming it *validation ratio* and using it to validate clinical treatment.

MacKinnon (1994) also used the ratio, referring to it as *proportion mediated* (p.137), for prevention intervention research. Barreto and Ellemers (2005) applied the ratio to study sexism and gender inequality. Ao et al. (2023) and Zhu et al. (2025) were among the latest to further develop and apply the ratio, which they call *percent contribution* (c_p), to study the impact of social media or e-health use on health.

Two simulation studies (MacKinnon et al. 1995; Tofighi et al. 2009) found the ratio unstable with small sample sizes, where "small" means 50, 100, 200, 500, 1,000 or 2,000, depending on effect sizes and other factors. Preacher and Kelley (2011) critiqued this ratio among other effect size indices, before introducing a new index, *maximum possible indirect effect* (k^2).

S3.2 Percentage Contribution t_p

Alwin and Hauser (1975) intended the ratio ab/c for models without counteraction, aka competition or suppression (Note 3, p. 43). In these *directionally complementary (di-plementary)* models, indirect path (ab) or the direct-and-remainder path (d) does not exceed the total effect (c) in strength ($|ab| < |c|$ and $|d| < |c|$), therefore the ab/c ratio stays within $0 \sim 100\%$ ($0 < ab/c < 1$, and $0 < d/c < 1$). The word "proportion" is a proper descriptor for the ratio in di-plementary models.

However, one of the real-data examples we report is *directionally competitive (di-petitive)*, where $|ab| > |c|$ and $|d| > |c|$, hence ab/c ratio ranges between $-\infty$ and ∞ excluding the range $0 \sim 1$. That is, let $t_{pi} = ab/c$, $t_{pi} < 0$ or $t_{pi} > 1$ for $abd < 0$. A proportion does not go below 0% or above 100% according to the common understanding of the English word "proportion" (Cambridge Dictionary 2025). "Proportion" is not an accurate descriptor of t_p in directionally competitive models.

In contrast, "percent" or "percentage" can exceed 100%, e.g., an increase of 200 percent means doubling (Cambridge Dictionary 2025). They can also be negative, e.g., a -50 percent means a reduction of half of the base amount. We therefore consider the ratio as a percentage, as MacKinnon and Dwyer (1993) did, or as percent, as Ao et al. (2023), and Zhu et al. (2025) did. The term "percentage" is chosen over "percent" as the general descriptor of the concept, because "percentage" is more of a general term while "percent" is expected to be preceded by a particular number, such as 50 percent (JAMA Network Editors 2020).

Accordingly, we will refer to the ratio as *Percentage contribution to total effect* and denote it as t_p for general use, covering the contribution of both indirect and direct-and-remainder paths, and possibly direct paths in more complex models. In a one-mediator simple mediation model, we will use t_{pi} to represent the percentage contribution of the indirect path (Eq. (S4)) and t_{pd} to represent that of the direct-and-remainder path (Eq. (S5)).

$$t_{pi} = \frac{a \cdot b}{c}, \tag{S4}$$

$$t_{pd} = \frac{d}{c}, \tag{S5}$$

where
t_{pd} =percentage contribution to total effect by indirect effect,
t_{pi} =percentage contribution to total effect by direct-and-remainder effect,
a =regression coefficient for the first leg of the mediated path,
b =regression coefficient for the second leg of the mediated path,
d =regression coefficient for the direct-and-remainder path,
c =regression coefficient for the total effect,
$t_p = t_{pi}$ or $t_p = t_{pd}$.

S3.3 Directional Typology of Models Involving Mediation (MIM)

S3.3.1 P-based Typology and P&D-based Typology

The distinction between *proportion mediated* and *percentage contribution* suggests the need to differentiate two types of *models involving mediation* (MIM), those with $abd > 0$, and those with $abd < 0$, regardless of p values or statistical acknowledgeability. It further implies the need for a supplemental system of MIM typology.

Two systems of typology are available. Baron and Kenny (1986) devised a typology of three types, i.e., complete (full) mediation, partial mediation, and no mediation. which continues to be highly influential in many disciplines. It is a *p-based typology. P* values, which indicate the level of statistical significance, serve as the partition lines between types. Zhao et al. (2010) used two criteria, p values and the direction of effect indicated by regression coefficient, to be the partition lines for a typology that includes three types of mediation and two types of non-mediation. This is a *p&direction-based*, or *p&d-based, typology.*

Both systems rely on p values as partition lines. To fully understand the behavior of percentage contribution (t_p), aka proportion mediated, in models involving mediation, we need another typology that uses the direction of effect as the only criteria of demarcation.

S3.3.2 Directional Typology of Models Involving Mediation (MIM)

Table S1 shows a typology of models involving mediation (MIM). These types are inclusive and mutually exclusive. This is a *directional typology*, or *d-based typology*, as the direction of the regression coefficient is the only criterion for partitioning.

While seven types were included, five (Lines $3 \sim 7$) are expected to be rare, especially for large samples. That's because each of the five types requires at least one parameter, e.g., ab, d, or c, to be exactly zero, which rarely occurs. Data analysts may expect to regularly encounter the other two types, directional complement (di-plement) and directional competition (di-petition), as shown in Lines 1 and 2 of Table S1.

Unlike the other two typologies, the directional typology is not designed to help establish mediation. That is, the typology does not attempt to identify the models of established mediation, although it does differentiate models of potential mediation from classified non-mediation (Column H, Table S1). The d-based typology is to supplement the p&d-based typology, to help better understand the parts, process, and product of models that may involve mediation (MIM).

S3.4 Differential Functions of Percentage Contribution (t_p)

S3.4.1 Two Functions of Statistical Indices: Comprehension and Comparison

A statistical index has two general functions, which are to aid comprehension and to aid comparison (Zhao et al. 2022, 2024; Zhao and Zhang 2014). To comprehend, one needs to interpret a measured number against the unit of the measurement scale, e.g., interpreting a t_p score against unit 1, which represents the $0-100\%$ percentage scale. To compare, one needs to contrast one measured number against another measured number on the same scale, e.g., contrasting two t_p scores against each other. To be useful, an index needs to regularly perform one or both of the two general functions.

S3.4.2 t_p May Aid Comprehension & Comparison in Complementary Models

Alwin and Hauser (1975) designed the proportion mediated for models in which direct and indirect effects do not "counteract one another" ($ab \times d \geq 0$, Note 3, p. 43), regardless of p values. These are *directionally complementary (di-plementary)* models. In these models, t_p is a true proportional measure ranging between 0% and 100% ($0 \leq t_p \leq 1$). That means t_p would be stable, therefore may informatively and consistently perform both general functions.

In other words, in a di-plementary model, a t_p index may be interpreted against the $0 \sim 100\%$ percentage/proportional scale, aiding comprehension to perform the first function, and two or more t_p indices may be contrasted against each other, aiding comparison to perform the second function.

S3.4.3 t_p May Aid Comparison in Competitive Models

Instability of t_p occurs when direct and indirect effects counteract each other, regardless of p values. These are *directionally competitive (di-petitive)* models. In these models t_p indices are not proportional measures, as they range between positive and negative infinities ($-\infty \leq t_p \leq \infty$). That's because "the total effect is less than the sum of the absolute effects, and some components may be larger than the total effect" (Alwin and Hauser 1975, Note 3, p. 43). The difference of a major component effect (ab or d) over the total effect (c) could approach infinity of either direction ($-\infty \leq (ab - c) \leq \infty$, $-\infty \leq (d - c) \leq \infty$). The larger the difference, the less stable the t_p indices are when referenced against the $0 \sim 100\%$ percentage/proportional scale, and the less informative the t_p indices are aiding comprehension. The underlying cause is that, when $|c|$ is small due to competition and offsetting between ab and d, the ab/c and d/c ratios can approach positive or negative infinity, and tiny changes in c can produce drastic swings in ab/c and d/c. That can make the two ratios unstable and uninformative in reference to the $0-100\%$ percentage scale.

In other words, in a di-petitive model, a t_p index may be unstable when interpreted against the $0 \sim 100\%$ percentage/proportional scale, depending on the differential sizes between $|c|$ and $|ab|$, and between $|c|$ and $|d|$. The smaller is $|c|$ relative to $|ab|$ and $|d|$, the less stable is t_p, and the less effective it is to perform the first function of aiding comprehension.

However, even with di-petitive models, t_p indices can still perform the second function of aiding comparison, according to an extension of the analysis above. As said, the instability of t_p is attributed to ab/c and d/c being drastically affected by c. The drastic effect is consequential only when one interprets ab/c and d/c against a conceptually fixed scale, such as the $0 \sim 100\%$ percentage/proportional scale, for the first function of aiding comprehension. It would not be as consequential when one compares two or more t_p indices with each other, which is for the second function of aiding comparison. In this situation c would have proportionally equal effects on all ratios being compared (ab/c, d/c and possibly other t_p indices if there are more indirect paths), because c is the only denominator in each ratio. Thus, the relative sizes of t_p indices would be stable when compared with each other.

In other words, in a di-petitive model, t_p indices are stable when compared with each other. Thus, two or more t_p indices can perform the second function of aiding comparison in a di-petitive model.

S3.5 Two Real-Data Examples

The two examples we report in the text also illustrate the differential functions of percentage contribution (t_p) in different models, as discussed below.

S3.5.1 Roles of t_p in Competitive Models

Model 1 (Figure 10) presents a directionally-competitive complete (di-petitive-lete) mediation, which is a subtype of directionally competitive (di-petitive) mediation. As said, in di-petitive models, t_p for ab and d can be large or huge when c is small or tiny, i.e., equal or near zero. Figure 10 shows a real-data example of this phenomenon, which is t_p =118K% for ab, $t_p = -119$K% for d, and $b_p = 0.000014$ for c.

In situations like this, small changes in c may produce large swings of t_p for ab and d, which would make t_p appear unstable compared to the $0 - 100\%$ percentage scale. The huge t_p values (e.g., t_p =118K% for ab and $t_p = -119$K% for d) indicate a tiny c ($b_p = 0.000014$); other than that, the exact numerical value of t_p for ab or d is not very informative when referenced against the $0 \sim 100\%$ scale. For example, $t_p = 118,001\%$ vs $t_p = 118,559\%$ does not necessarily indicate a large difference despite the difference of almost 500%. That's the reason for rounding off at a high level like 1,000, e.g., 118K% and −119K%.

Despite the low ability to aid comprehension, however, t_p indices can still be informatively compared with each other, thereby aiding comparison. Comparing t_p =118K% for ab with $t_p = -119$K% for d, for example, one may obtain important information, such as: 1) The opposing signs of the two t_p indices indicate a directionally competitive relation, even though the d path fails to pass the statistical test ($p = .6411$). 2) The direct and indirect paths are about equal in strengths (118K% vs −119K%), indicating approximately even competition. 3) The about even competition explains the tiny total effect ($b_p = 0.000014$ for c).

S3.5.2 Roles of t_p in Complementary Models

Model 2 (Figure 11) presents a directionally-complementary complete (di-plementary-lete) mediation, which is a subtype of directionally complementary (di-plementary) mediation. As seen in Figure 11, the total effect is positive ($b_p = .0243$, $t_p = 100\%$, $p = .2704$). Of the total effect, indirect path contributes a little below 30% ($t_p \approx 29\%$) and the direct-and-remainder (di-remainder) path contributes a little over 70% ($t_p \approx 71\%$). Notably, the path that contributes comparatively more (71%, di-remainder path) failed the statistical test ($p = .2704$), while the path that contributes comparatively less (29%, indirect path), passed the statistical test with margin ($p < .001$). In our experience, this has been often-seen phenomenon with this subtype of mediation, i.e., directionally complementary complete (di-plementary-lete) mediation.

The analysis above utilizes 1) the comprehension function of t_p, which requires to reference the t_p scores, namely $t_p = 29\%$ and $t_p = 71\%$, against the percentage scale, and 2) the comparison function of t_p, which requires to contrast the t_p scores, $t_p = 29\%$ and $t_p = 71\%$ against each other. Alwin and Hauser (1975) intended the proportion mediated for this type of mediation, i.e., the di-plementary mediation (p. 43, Note 3). It is not surprising that t_p, which is an extension of proportion mediated, can informatively and effectively perform both functions to aid comprehension and comparison.

S3.6 Conclusion

The unstableness of percentage contribution (t_p), aka proportion mediated, occurs in directionally competitive models, where $abd < 0$. Larger differences between $|ab|$ and $|c|$ ($|ab| - |c|$) produce higher unstableness, assuming $abd < 0$, hence $|ab| > |c|$. Unstableness affects comprehensibility, but not comparability, of t_p scores. Therefore, in directionally competitive models, percentage contribution (t_p) may aid comparison more effectively than it may aid interpretation. That is, in a competitive model, 1) comparing two t_p scores would be informative, but 2) interpreting a t_p score against the $0 \sim 100\%$ percentage scale would be less informative, especially when the difference between $|ab|$ and $|c|$ is large, therefore the ratio $|t_{pi}| = |ab/c|$ is far above 1 (100%).

The unstableness of t_p is not expected to occur in directionally complementary models, where $abd > 0$. In these models, percentage contribution (t_p) may aid comprehension and comparison both effectively.

Table S1 Directional Typology of Models Involving Mediation and Ranges of Percent Contribution to Total Effect (t_p)

A	B	C	D	E	F	G	H
	Directional Mediation Model	Short Name	Model Definition	t_{pi}–t_p of indirect path	t_{pd}–t_p of direct path	Theoretical range of t_p	Potential Mediation
1	directional complement	di-plement	$abd > 0$ & $c \neq 0$	ab/c	d/c	$0 < t_p < 1$	Yes
2	directional competition	di-petition	$abd < 0$ & $c \neq 0$	ab/c	d/c	$-\infty < t_p < 0$ or $1 < t_p < \infty$	Yes
3	directional completion	di-pletion	$d = 0$ & $c \neq 0$	ab/c	d/c	$t_{pi} = 1$ & $t_{pd} = 0$	Yes
4	directional non-mediation	non-mediation	$ab = 0$ & $c \neq 0$	ab/c	d/c	$t_{pi} = 0$ & $t_{pd} = 1$	No
5	even competition & positive mediation	e-comp$^+$	$ab > 0$ & $c = 0$	∞	$-\infty$	$t_{pi} = \infty$ & $t_{pd} = -\infty$	Yes
6	even competition & negative mediation	e-comp$^-$	$ab < 0$ & $c = 0$	$-\infty$	∞	$t_{pi} = -\infty$ & $t_{pd} = \infty$	Yes
7	zero indirect and direct effects	zero effects	$ab = 0$ & $c = 0$	0	0	$t_{pi} = 0$ & $t_{pd} = 0$	No

References

Alwin DF, Hauser RM (1975) The decomposition of effects in path analysis. American sociological review pp 37–47

Ao SH, Zhang L, Liu PL, et al (2023) Social media and partnership jointly alleviate caregivers' psychological distress: exploring the effects of online and offline connectedness. BMC psychology 11(1):394

Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 51(6):1173–1182

Barreto M, Ellemers N (2005) The burden of benevolent sexism: How it contributes to the maintenance of gender inequalities. European journal of social psychology 35(5):633–642

Cambridge Dictionary (2025) Proportion. URL https://dictionary.cambridge.org/dictionary/english/proportion, retrieved 5 April 2025

Ditlevsen S, Christensen U, Lynch J, et al (2005) The mediation proportion: a structural equation approach for estimating the proportion of exposure effect on outcome explained by an intermediate variable. Epidemiology 16(1):114–120

Freedman LS (2001) Confidence intervals and statistical power of the 'validation'ratio for surrogate or intermediate endpoints. Journal of Statistical Planning and Inference 96(1):143–153

Freedman LS, Graubard BI, Schatzkin A (1992) Statistical validation of intermediate endpoints for chronic diseases. Statistics in medicine 11(2):167–178

Huang B, Sivaganesan S, Succop P, et al (2004) Statistical assessment of mediational effects for logistic mediational models. Statistics in medicine 23(17):2713–2728

JAMA Network Editors (2020) AMA Manual of Style: A Guide for Authors and Editors, 11th edn. Oxford University Press

Jiang Y, Zhao X, Zhu L, et al (2021) Total-effect test is superfluous for establishing complementary mediation. Statistica Sinica 31(4):1961–1983

MacKinnon DP (1994) Analysis of mediating variables in prevention intervention studies. In: Cazares A, Beatty LA (eds) Scientific Methods for Prevention Intervention Research, vol 139. U.S. Department of Health and Human Services, p 127–153

MacKinnon DP, Dwyer JH (1993) Estimating mediated effects in prevention studies. Evaluation review 17(2):144–158

MacKinnon DP, Warsi G, Dwyer JH (1995) A simulation study of mediated effect measures. Multivariate behavioral research 30(1):41–62

Preacher KJ, Kelley K (2011) Effect size measures for mediation models: quantitative strategies for communicating indirect effects. Psychological methods 16(2):93

Sobel ME (1982) Asymptotic confidence intervals for indirect effects in structural equation models. Sociological methodology 13:290–312

Tofighi D, MacKinnon DP, Yoon M (2009) Covariances between regression coefficient estimates in a single mediator model. British Journal of Mathematical and Statistical Psychology 62(3):457–484

Zhao X, Zhang XJ (2014) Emerging methodological issues in quantitative communication research. New trends in communication studies, II pp 953–978

Zhao X, Lynch Jr JG, Chen Q (2010) Reconsidering baron and kenny: Myths and truths about mediation analysis. Journal of consumer research 37(2):197–206

Zhao X, Feng GC, Ao SH, et al (2022) Interrater reliability estimators tested against true interrater reliabilities. BMC Medical Research Methodology 22(1):232

Zhao X, Li DM, Lai ZZ, et al (2024) Percentage coefficient (bp)–effect size analysis (theory paper 1). arXiv preprint arXiv:240419495

Zhu Y, Xiao QE, Ao MC, et al (2025) How ehealth use and cancer information-seeking influence older adults' acceptance of genetic testing: Mediating roles of pigi and cancer worry. Digital Health 11:20552076251317658