JOINT SIGNIFICANCE TESTS FOR MEDIATION EFFECTS OF SOCIOECONOMIC ADVERSITY ON ADIPOSITY VIA EPIGENETICS[1]

BY YEN-TSUNG HUANG

Academia Sinica

Mediation analysis has become a popular practice in biomedical research. We conduct mediation analyses to investigate whether epigenetic variations mediate the effect of socioeconomic disadvantage on adiposity. Mediation effects can be expressed as a product of two parameters: one for the exposure-mediator association and the other for the mediator-outcome association conditional on the exposure. Under multi-mediator models, we study joint significance tests which examine the two parameters separately and compare with the widely used product significance tests which focus on the product of two parameters. Normal approximation of product significance tests depends on both effect size and sample size. We show that joint significance tests are intersection-union tests with size α and asymptotically more powerful than the normality-based product significance tests. Based on the theoretical results, we construct powerful testing procedures for gene-based mediation analyses and path-specific analyses. Advantage of joint significance tests is supported by simulation as well as the results of locus-based and gene-based mediation analyses of chromosome 17. Our analyses suggest that methylation of *FASN* gene mediates the effect of socioeconomic adversity on adiposity.

1. Introduction. Mediation analysis first proposed in psychological literature has been a popular approach [Baron and Kenny (1986), MacKinnon (2008)]. As illustrated in Figure 1a with a directed acyclic graph [Robins (2003)], the mediation model includes an exposure S, a mediator M, and an outcome Y. Mediation analysis decomposes the effect of the exposure on the outcome into an indirect effect mediated through the mediator and a direct effect not through the mediator, aiming for better understanding of the underlying mechanism. By employing the counterfactuals [Rubin (1978)], ignorability assumptions for effect identifiability have been carefully studied [Imai, Keele and Yamamoto (2010), Pearl (2001), Robins and Greenland (1992), VanderWeele and Vansteelandt (2009)]. Built upon that, advanced methodology of mediation analysis has been developed for dichotomous outcomes [VanderWeele and Vansteelandt (2010)] as well as time-to-event survival

(a) Single-mediator model   (b) Multi-mediator model without ordering   (c) Two-mediator model with ordering
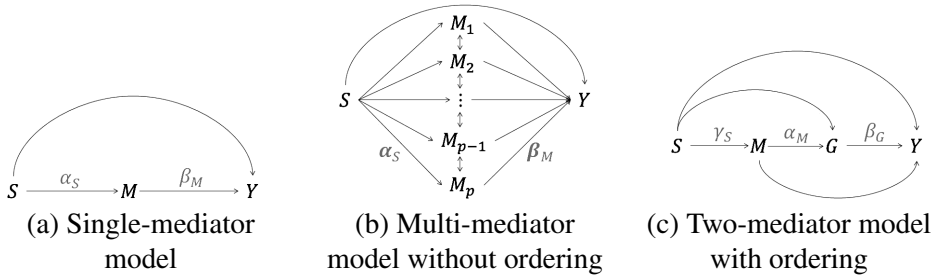
FIG. 1. *Directed acyclic graph (DAG) of causal mediation models.*

outcomes [Lange and Hansen (2011), Tchetgen Tchetgen (2011), VanderWeele (2011)].

This paper is motivated by a study where 285,163 epigenetic DNA methylation loci were investigated one at a time as a potential mediator in relation to the effect of socioeconomic adversity on adiposity measured by body mass index (BMI) [Loucks et al. (2016)]. Epigenetics has been known as an "epicenter" that integrates the influence of environment and genetics [Loucks et al. (2016)]. Childhood socioeconomic disadvantage is associated with obesity [Senese et al. (2009)] and epigenetic DNA methylation [Borghol et al. (2012)]. Genome-wide analysis has shown that epigenetic marks are associated with later life obesity risk in New England Family Study (NEFS) [Agha et al. (2015)]. Based on their pairwise relationships in existing literature, we hypothesize that adiposity affected by socioeconomic status may arise as results of alterations of biological mechanisms via epigenetic regulation.

Suppose that the mediation model as illustrated in Figure 1a contains BMI and the mediator of DNA methylation level. Under this model, we further assume that BMI Y is linearly determined by the socioeconomic index S and the mediator M (DNA methylation): $E(Y) = \beta_0 + \beta_S S + \beta_M M$, and the mediator is linearly determined by the exposure S: $E(M) = \alpha_0 + \alpha_S S$. The mediation effect can be expressed as a product of the two parameters, $\alpha_S \beta_M$ [Baron and Kenny (1986), MacKinnon (2008)]. The product expression for the mediation effect is shared by mediation analyses for dichotomous or survival outcome under certain model specification and can be easily extended to incorporate the exposure-by-mediator interaction [Lange and Hansen (2011), VanderWeele (2011), VanderWeele and Vansteelandt (2010)]. Two classes of hypothesis tests have been proposed to test $H_0 : \alpha_S \beta_M = 0$. One is to estimate α_S and β_M with maximum likelihood estimators $\hat{\alpha}_S$ and $\hat{\beta}_M$, respectively, and to compare $\hat{\alpha}_S \hat{\beta}_M$ with its underlying distribution, termed product significance test (PT). The underlying distribution can be approximated by Gaussian distribution using delta method [Sobel (1982)], normal product distribution (NP), that is, the distribution for product of two normals [MacKinnon et al. (2002)] or bootstrapping [Bollen and Stine (1990), Preacher and Hayes (2008)]. The other class of tests examines $H_0 : \alpha_S = 0$ and $H_0 : \beta_M = 0$

separately, termed joint significance test (JT) [MacKinnon et al. (2002)], and JT is statistically significant if both tests are significant.

Simulation studies by MacKinnon et al. (2002) suggested that PT with bootstrapping had better power than PT with various distribution approximations. For the motivating epigenomic study, PT with bootstrapping is not feasible because of its computation cost. We show in data application that analyses of DNA methylation loci in chromosome 17 with 1000 bootstrap replicates take 16 hours, and the projected computation time for all chromosomes is >11 days with 1000 replicates and >300 years with precision to reach the genome-wide significance level. Therefore, there is an imperative need for analytic approaches. The simulation of MacKinnon et al. (2002) also suggested that JT may have decent power over PT, which had a lack of theoretical justification. Furthermore, little is studied regarding validity and test size of JT. Besides single-locus analyses, we were particularly interested to know whether all methylation loci within a gene *en bloc* mediate the socioeconomic effect on BMI, which is the so-called gene-based analyses. Gene-based approach in genome-wide analyses has been shown superior in statistical power and biological relevance [Wu et al. (2010)]. To our knowledge, however, none can be found for JT about its application or generalization to multi-mediator models. This paper aims to characterize the asymptotic properties of JT and PT; moreover, bootstrap-based PT (PT-B) and a recently developed method based on normal product distribution (PT-NP) [Huang and Pan (2016)] are compared via simulation.

The paper is structured as follows. In Section 2, we introduce the New England Family Study. In Section 3, we study multivariate mediation effects. In Section 4, we show that joint significance tests are intersection-union tests with size α. In Section 5, we conduct extensive simulation studies to evaluate the theoretical results. The epigenetic study is presented in Section 6. The paper concludes with a discussion in Section 7.

2. The New England Family Study (NEFS).   The data used for mediation analyses were nested within the New England Family Study [Huang et al. (2016)], which is comprised of 17,921 offspring of pregnant women in the Collaborative Perinatal Project from Providence (Rhode Island) and Boston (Massachusetts) in the United States (US), recruited between 1959 and 1974 [Hardy (1971)]. We conducted mediation analyses in 74 women with adequate adipose tissue collected from needle biopsy for genome-wide epigenetic profiling. Childhood socioeconomic disadvantage at age 7 was assessed by a socioeconomic index, which summarized average percentile of both parents' educational attainment, occupation, and income relative to the US population [Loucks et al. (2016)]; and adiposity was directly assessed using BMI (kg/m^2) at mean age 47 years. DNA methylation of biopsy adipose tissue samples collected at adulthood was evaluated using the Infinium HumanMethylation450K BeadChip (Illumina, San Diego, CA). We performed both locus-centric (Figure 1a) and gene-centric analyses (Figure 1b),

examining 16,394 loci and 1151 genes on chromosome 17. There were 24 methylation loci and 9 genes with $p < 0.05$ using normality-based product significance test (PT-N). We are also interested in the mediation effect by epigenetics that is further mediated by the childhood BMI. It is the so-called path-specific effects (PSEs), that is, mediation effects through certain sequences of mediators [Avin, Shpitser and Pearl (2005), Taylor, MacKinnon and Tein (2008), VanderWeele and Vansteelandt (2013)]. To this end, we set up the mediation model in Figure 1c, with S, M, G, and Y being socioeconomic index, BMI at age 7, DNA methylation, and BMI at adulthood, respectively. Normality-based tests for PSE through epigenetics of an obesity-related genes *FASN* and childhood BMI showed no promising results [Table S1 in the Supplementary Material Huang (2018)].

The study demonstrates the utility of mediation analyses in interrogating the mechanism of well-established exposure-outcome relationships. However, we also found that the conventional mediation test was not a powerful test, and the normal approximation may not hold under certain circumstances. To address these, we propose two multivariate JTs: one for the marginal mediation effect with multiple mediators (Figure 1b), and the other one for path-specific effects (Figure 1c).

3. Hypothesis test of mediation effect. We propose two joint significance tests and rigorously study their theoretical properties prior to applying to the epigenetic studies.

3.1. *Hypothesis test of* $\mathbf{U}_n^T \mathbf{V}_n$. Consider a mediation model in Figure 1b that can be expressed with two models:

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta}_X + S_i \beta_S + \mathbf{M}_i^T \boldsymbol{\beta}_M + \varepsilon_{Yi}, \tag{1}$$

$$\mathbf{M}_i = \mathbf{A}_X \mathbf{X}_i + S_i \boldsymbol{\alpha}_S + \boldsymbol{\varepsilon}_{Mi}, \tag{2}$$

where Y_i, $\mathbf{M}_i = (M_{1i}, \ldots, M_{pi})^T$, S_i and $\mathbf{X}_i$ are the outcome BMI, the mediators DNA methylation levels, the exposure measured by a socioeconomic index and covariates, respectively, for subject i, $i = 1, \ldots, n$, $\mathbf{X}_i$ is a q-by-1 vector with the first element being 1 (the intercept), $\mathbf{A}_X$ is a p-by-q matrix, $\boldsymbol{\beta}_M = (\beta_{M1}, \ldots, \beta_{Mp})^T$, $\boldsymbol{\alpha}_S = (\alpha_{S1}, \ldots, \alpha_{Sp})^T$, $\varepsilon_Y \sim N(0, \sigma_Y^2)$, $\boldsymbol{\varepsilon}_M \sim N(\mathbf{0}, \Sigma_M)$, and Σ_M is a p-by-p bounded covariance matrix. Under assumptions of no unmeasured confounding, ε_Y and $\boldsymbol{\varepsilon}_M$ are independent and do not depend on $\boldsymbol{\alpha}_S$ or $\boldsymbol{\beta}_M$, and the mediation effect can be expressed as $\boldsymbol{\alpha}_S^T \boldsymbol{\beta}_M$ [Huang and Pan (2016)].

Model (1) can be extended to incorporate logistic regression models [VanderWeele and Vansteelandt (2010)], additive hazard models [Lange and Hansen (2011)], Cox proportional hazard models [VanderWeele (2011)], with the following general model: $h_i = \mathbf{X}_i^T \boldsymbol{\beta}_X + S_i \beta_S + \mathbf{M}_i^T \boldsymbol{\beta}_M$, where $h_i = \text{logit} \Pr(Y_i = 1)$ for logistic models, $h_i = \lambda_T(t \mid \mathbf{X}_i, S_i, \mathbf{M}_i)$, the hazard, for additive hazard models (T is survival time), and $h_i = \log \lambda_T(t \mid \mathbf{X}_i, S_i, \mathbf{M}_i)$ for Cox models. Other extensions such as accelerated failure time (AFT) models [VanderWeele (2011)]

and semiparametric probit models [Huang and Cai (2016)] can be expressed as: $h(T_i) = \mathbf{X}_i^T \boldsymbol{\beta}_X + S_i \beta_S + \mathbf{M}_i^T \boldsymbol{\beta}_M + \varepsilon_{Yi}^*$, where $\varepsilon_Y^* \sim N(0, 1)$ in probit models and follows a parametric distribution such as Weibull or a nonparametric distribution in AFT models. The expression of mediation effects from these extensions is identical to $\boldsymbol{\alpha}_S^T \boldsymbol{\beta}_M$, with interpretation depending on the effect scale and model assumptions. Under these models, maximum likelihood estimators of $\boldsymbol{\alpha}_S$ and $\boldsymbol{\beta}_M$ scaled by $\sqrt{n}$ are asymptotically normal, that is, $\sqrt{n}(\hat{\boldsymbol{\alpha}}_S - \boldsymbol{\alpha}_S) \to N(\mathbf{0}, \Sigma_\alpha)$ and $\sqrt{n}(\hat{\boldsymbol{\beta}}_M - \boldsymbol{\beta}_M) \to N(\mathbf{0}, \Sigma_\beta)$ where Σ_α and Σ_β are bounded matrices. Another extension is to incorporate S-by-$\mathbf{M}$ interaction in model (1) by replacing $\boldsymbol{\beta}_M$ with $\boldsymbol{\beta}_M + \boldsymbol{\beta}_{SM} S$. The scaled estimator $\sqrt{n}(\hat{\boldsymbol{\beta}}_M + \hat{\boldsymbol{\beta}}_{SM} S)$ still follows Gaussian distribution asymptotically.

Since mediation effects from the above models share the same expression, a product of the two parameters $\boldsymbol{\alpha}_S$ and $\boldsymbol{\beta}_M$, we focus on the null hypothesis: $H_0 : \boldsymbol{\alpha}_S^T \boldsymbol{\beta}_M = 0$. Denote $\boldsymbol{\theta}^* = (\theta_1^*, \ldots, \theta_p^*)^T$ and θ_j^* is either α_{Sj} or β_{Mj}. Here, we do not consider perfect cancellation of component-wise mediation effects, that is, at least one of α_{Sj} and β_{Mj} is 0 for all j under the null, and such a null hypothesis is equivalent to the following null:

$$H_0 : \bigcup_{k=1,\ldots,2^p} \Theta_{\mathbf{U}^T \mathbf{V}, k},$$

(3)

$$\Theta_{\mathbf{U}^T \mathbf{V}, k} = \left\{ \boldsymbol{\theta} = (\boldsymbol{\alpha}_S^T, \boldsymbol{\beta}_M^T)^T : \sum_{j=1}^p w_j \theta_j^* = 0, \theta_j^* \in \{\alpha_{Sj}, \beta_{Mj}\}, \boldsymbol{\theta} \setminus \boldsymbol{\theta}^* \in \Re^p \right\},$$

where w_j is an arbitrary nonzero weight. We compare the asymptotic product significance test based on $n\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M$ and a joint significance test proposed in the following.

LEMMA 3.1. *Suppose that $\mathbf{U}_n$ and $\mathbf{V}_n$ are two independent multivariate Gaussian variables with respective means $\boldsymbol{\mu}_{1n}$, $\boldsymbol{\mu}_{2n}$ and bounded covariances Σ_1, Σ_2, where $\mathbf{U}_n$ and $\mathbf{V}_n$ are sequences of random variables, and $\boldsymbol{\mu}_{1n}$ and $\boldsymbol{\mu}_{2n}$ are sequences of vectors indexed by n. If $\boldsymbol{\mu}_{1n}^T \Sigma_2 \boldsymbol{\mu}_{1n} + \boldsymbol{\mu}_{2n}^T \Sigma_1 \boldsymbol{\mu}_{2n} \to \infty$ as $n \to \infty$, then $\frac{\mathbf{U}_n^T \mathbf{V}_n - \boldsymbol{\mu}_{1n}^T \boldsymbol{\mu}_{2n}}{\sqrt{\boldsymbol{\mu}_{1n}^T \Sigma_2 \boldsymbol{\mu}_{1n} + \boldsymbol{\mu}_{2n}^T \Sigma_1 \boldsymbol{\mu}_{2n}}} \to N(0, 1)$.*

REMARK 1. Letting $\mathbf{U}_n \equiv \sqrt{n}\hat{\boldsymbol{\alpha}}_S$ and $\mathbf{V}_n \equiv \sqrt{n}\hat{\boldsymbol{\beta}}_M$, one can use Lemma 3.1 to show that $\frac{n(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M - \boldsymbol{\alpha}_S^T \boldsymbol{\beta}_M)}{\sqrt{n\boldsymbol{\alpha}_S^T \Sigma_\beta \boldsymbol{\alpha}_S + n\boldsymbol{\beta}_M^T \Sigma_\alpha \boldsymbol{\beta}_M}}$ converges to standard normal distribution, as $n \to \infty$. To satisfy the condition $\boldsymbol{\mu}_{1n}^T \Sigma_2 \boldsymbol{\mu}_{1n} + \boldsymbol{\mu}_{2n}^T \Sigma_1 \boldsymbol{\mu}_{2n} = n\boldsymbol{\alpha}_S^T \Sigma_\beta \boldsymbol{\alpha}_S + n\boldsymbol{\beta}_S^T \Sigma_\alpha \boldsymbol{\beta}_S \to \infty$, one needs either large effects or sample size under the alternative hypothesis, and at least one large nonzero effect and large sample size under the null hypothesis. We denote two test statistics $T_{\mathbf{U}^T \mathbf{V}} = \frac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M - 0)^2}{\boldsymbol{\alpha}_S^T \Sigma_{\beta n} \boldsymbol{\alpha}_S + \boldsymbol{\beta}_M^T \Sigma_{\alpha n} \boldsymbol{\beta}_M}$ for PT-N*

and $\hat{T}_{\mathbf{U}^T\mathbf{V}} = \frac{(\hat{\boldsymbol{\alpha}}_S^T\hat{\boldsymbol{\beta}}_M - 0)^2}{\hat{\boldsymbol{\alpha}}_S^T\Sigma_{\beta n}\hat{\boldsymbol{\alpha}}_S + \hat{\boldsymbol{\beta}}_M^T\Sigma_{\alpha n}\hat{\boldsymbol{\beta}}_M}$ for PT-N where $n\Sigma_{\beta n} \to \Sigma_\beta$ and $n\Sigma_{\alpha n} \to \Sigma_\alpha$ as $n \to \infty$.

With $p = 1$, Lemma 3.1 is simplified to an alternative version of the theorem in Aroian (1947) stating that the product of two normal variables standardized by the product of their standard deviation, σ_{1n} and σ_{2n} converges to normal as $\frac{\mu_{1n}}{\sigma_{1n}}$ or $\frac{\mu_{2n}}{\sigma_{2n}}$ goes to infinity. We next propose a joint significance test for null (3).

DEFINITION 1 (JT of $\mathbf{U}_n^T\mathbf{V}_n$). Let $\Sigma_n = \begin{bmatrix} \Sigma_{\alpha n} & 0 \\ 0 & \Sigma_{\beta n} \end{bmatrix}$, $\mathbf{w} = (w_1, \ldots, w_p)^T$, $w_j \in \{\hat{\alpha}_{Sj}, \hat{\beta}_{Mj}\}$, $\mathbf{w}^* = (w_1^*, \ldots, w_p^*, w_{p+1}^*, \ldots, w_{2p}^*)^T$ where if $w_j = \hat{\alpha}_{Sj}$, then $w_j^* = 0$ and $w_{p+j}^* = \hat{\alpha}_{Sj}$; if $w_j = \hat{\beta}_{Mj}$, then $w_j^* = \hat{\beta}_{Mj}$ and $w_{p+j}^* = 0$. Calculate $\{\mathbf{w}^{*T}\Sigma_n\mathbf{w}^*\}$ over 2^p possible combinations of w_j and perform a two-sided z-test on $\frac{|\hat{\boldsymbol{\alpha}}_S^T\hat{\boldsymbol{\beta}}_M|}{\sqrt{V_{\max}}}$, where $V_{\max} = \max\{\mathbf{w}^{*T}\Sigma_n\mathbf{w}^*\}$.

It is trivial to show that with $p = 1$, the above procedure is simplified to the univariate JT where one tests $H_0 : \alpha_S = 0$ and $H_0 : \beta_M = 0$, and picks the larger p-value [MacKinnon et al. (2002)]. The procedure can be construed as a multivariate generalization by weighting the parameter (e.g., the mediator-outcome association) with the other (e.g., the exposure-mediator association). Asymptotic independence of $\sqrt{n}\hat{\boldsymbol{\alpha}}_S$ and $\sqrt{n}\hat{\boldsymbol{\beta}}_M$ can be established by the second derivative of the joint log-likelihood of models (1) and (2): $\frac{\partial^2 \log f(Y,\mathbf{M}|\mathbf{X},S)}{\partial\boldsymbol{\beta}_M \partial\boldsymbol{\alpha}_S^T} = \frac{\partial^2 \log f(Y|\mathbf{X},S,\mathbf{M})}{\partial\boldsymbol{\beta}_M \partial\boldsymbol{\alpha}_S^T} + \frac{\partial^2 \log f(\mathbf{M}|\mathbf{X},S)}{\partial\boldsymbol{\beta}_M \partial\boldsymbol{\alpha}_S^T} = \mathbf{0}$, where $f(Y \mid \mathbf{X}, S, \mathbf{M})$ and $f(\mathbf{M} \mid \mathbf{X}, S)$ are probability density functions of models (1) and (2), respectively. We note that $\mathbf{w}$ in Definition 1 can be specified as any arbitrary weights, for example, $\mathbf{w} = \mathbf{1}$. We choose the above weighting scheme such that we are able to compare its performance with the product significance tests. The test conditional on the estimated weights can be viewed as a test conditional on the data estimating the weighting parameters.

THEOREM 3.1. *Suppose that* $\frac{\boldsymbol{\alpha}_S^T\Sigma_{\beta n}\boldsymbol{\alpha}_S + \boldsymbol{\beta}_M^T\Sigma_{\alpha n}\boldsymbol{\beta}_M}{\hat{\boldsymbol{\alpha}}_S^T\Sigma_{\beta n}\hat{\boldsymbol{\alpha}}_S + \hat{\boldsymbol{\beta}}_M^T\Sigma_{\alpha n}\hat{\boldsymbol{\beta}}_M} \to 1$ *and* $n^2\boldsymbol{\alpha}_S^T\Sigma_{\beta n}\boldsymbol{\alpha}_S +$ $n^2\boldsymbol{\beta}_M^T\Sigma_{\alpha n}\boldsymbol{\beta}_M \to \infty$ *as* $n \to \infty$, *and* $\alpha_{Sj}\alpha_{Sk}\rho_{\beta,jk} \geq 0$ *and* $\beta_{Mj}\beta_{Mk}\rho_{\alpha,jk} \geq 0$, *where* $\rho_{\alpha,jk}$ *is covariance of* $\sqrt{n}\hat{\alpha}_{Sj}$ *and* $\sqrt{n}\hat{\alpha}_{Sk}$ *and* $\rho_{\beta,jk}$ *is that of* $\sqrt{n}\hat{\beta}_{Mj}$ *and* $\sqrt{n}\hat{\beta}_{Mk}$. *Under the null* (3), *the JT in Definition* 1 *is asymptotically more powerful than the PT based on* $\hat{T}_{\mathbf{U}^T\mathbf{V}}$.

The condition $n^2\boldsymbol{\alpha}_S^T\Sigma_{\beta n}\boldsymbol{\alpha}_S + n^2\boldsymbol{\beta}_M^T\Sigma_{\alpha n}\boldsymbol{\beta}_M \to \infty$ (e.g., large sample size or effect) is for normal approximation of PT. JT works regardless of the effect size. What Theorem 3.1 establishes is that *even under the condition that normal approximation works well for PT*, it is still less powerful than JT. The intuition is

that under the null of $\boldsymbol{\alpha}_S \neq \mathbf{0}$ and $\boldsymbol{\beta}_M = \mathbf{0}$, for example, $\frac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{\hat{\boldsymbol{\alpha}}_S^T \Sigma_{\beta n} \hat{\boldsymbol{\alpha}}_S + \hat{\boldsymbol{\beta}}_M^T \Sigma_{\alpha n} \hat{\boldsymbol{\beta}}_M} = \frac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{V_{\max}} (1 - \frac{\hat{\boldsymbol{\beta}}_M^T \Sigma_{\alpha n} \hat{\boldsymbol{\beta}}_M}{\hat{\boldsymbol{\alpha}}_S^T \Sigma_{\beta n} \hat{\boldsymbol{\alpha}}_S + \hat{\boldsymbol{\beta}}_M^T \Sigma_{\alpha n} \hat{\boldsymbol{\beta}}_M}) = \frac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{V_{\max}} - c$, where $c = |o_p(1)|$ under H_0 and $|O_p(1)|$ under H_1 ($\boldsymbol{\beta}_M \neq \mathbf{0}$). Therefore, PT and JT follow the same distribution under the null, but the difference c guarantees a better power in JT. For implementation, one may transform the mediators to be uncorrelated conditional on S to ensure $\hat{\rho}_{\beta,jk} = \hat{\rho}_{\alpha,jk} = 0$, and thus $\alpha_{Sj} \alpha_{Sk} \rho_{\beta,jk} \approx \beta_{Mj} \beta_{Mk} \rho_{\alpha,jk} \approx 0$. Specifically, we orthogonally diagonalize the sample covariance of residuals in model (2) $\hat{\Sigma}_M$ by $\mathbf{u} \hat{\Sigma}_M \mathbf{u}^T = \mathrm{diag}(v_1, \ldots, v_p)$ and perform the proposed tests on the transformed mediators $\mathbf{P}_i = \mathbf{u} \mathbf{M}_i$ [Huang and Pan (2016)].

3.2. *Hypothesis test of $U_n V_n W_n$.* Tests for the mediation effect $\boldsymbol{\alpha}_S^T \boldsymbol{\beta}_M$ examine the marginal mediation effect without characterizing the effect mediated through specific mediators. To address this limitation, mediation effects through certain sequences of mediators or the path-specific effects (PSEs) have been proposed [Avin, Shpitser and Pearl (2005), Taylor, MacKinnon and Tein (2008), VanderWeele and Vansteelandt (2013)]. Here, we study PSEs in the model with two mediators (Figure 1c).

We propose the following three models to represent the two-mediator mediation model, as illustrated in Figure 1c:

$$ (4) \qquad Y_i = \mathbf{X}_i^T \boldsymbol{\beta}_X + S_i \beta_S + M_i \beta_M + G_i \beta_G + \varepsilon_{Yi}, $$

$$ (5) \qquad G_i = \mathbf{X}_i^T \boldsymbol{\alpha}_X + S_i \alpha_S + M_i \alpha_M + \varepsilon_{Gi}, $$

$$ (6) \qquad M_i = \mathbf{X}_i^T \boldsymbol{\gamma}_X + S_i \gamma_S + \varepsilon_{Mi}, $$

where $\varepsilon_Y \sim N(0, \sigma_Y^2)$, $\varepsilon_G \sim N(0, \sigma_G^2)$, and $\varepsilon_M \sim N(0, \sigma_M^2)$, all independent of β_G, α_M, γ_S. Under the assumption of no unmeasured confounding, ε_Y, ε_G, and ε_M are also independent, and the effect of socioeconomic adversity S on adult BMI Y mediated through childhood BMI M and possibly through DNA methylation G can be expressed as $\gamma_S(\beta_M + \alpha_M \beta_G)$. One can further identify the effect of S on Y mediated through M and G (the horizontal path in Figure 1c) with effect size of $\gamma_S \alpha_M \beta_G$ under strong identifiability assumptions [Albert and Nelson (2011), Taylor, MacKinnon and Tein (2008)] or the model assumption $\beta_M = 0$. The null hypothesis for this path-specific effect is

$$ (7) \qquad H_0 : \gamma_S \alpha_M \beta_G = 0. $$

Similar to Section 3.1, Y can be dichotomous or survival outcomes, and under certain model specification, the path-specific effect can be expressed as $\gamma_S \alpha_M \beta_G$. Suppose that $\hat{\gamma}_S$, $\hat{\alpha}_M$, and $\hat{\beta}_G$ are MLEs of γ_S, α_M, and β_G, respectively, and $\sqrt{n}(\hat{\gamma}_S - \gamma_S) \to N(0, \sigma_\gamma^2)$, $\sqrt{n}(\hat{\alpha}_M - \alpha_M) \to N(0, \sigma_\alpha^2)$, $\sqrt{n}(\hat{\beta}_G - \beta_G) \to N(0, \sigma_\beta^2)$ where σ_γ^2, σ_α^2 and σ_β^2 are all bounded. One can construct PT of null (7) by comparing $n^{3/2} \hat{\gamma}_S \hat{\alpha}_M \hat{\beta}_G$ with its underlying distribution.

LEMMA 3.2. *Suppose that U_n, V_n, and W_n follow independent Gaussian distributions with respective means μ_{1n}, μ_{2n}, and μ_{3n}, and bounded variances σ_1^2, σ_2^2, and σ_3^2, where U_n, V_n, W_n are sequences of random variables indexed by n. If at least two of μ_{1n}, μ_{2n}, μ_{3n} go to infinity as $n \to \infty$, then $\frac{U_n V_n W_n - \mu_{1n}\mu_{2n}\mu_{3n}}{\sqrt{\mu_{1n}^2\mu_{2n}^2\sigma_3^2 + \mu_{1n}^2\mu_{3n}^2\sigma_2^2 + \mu_{2n}^2\mu_{3n}^2\sigma_1^2}}$ converges to the standard normal.*

REMARK 2. Let $U_n \equiv \sqrt{n}\hat{\gamma}_S$, $V_n \equiv \sqrt{n}\hat{\alpha}_M$ and $W_n \equiv \sqrt{n}\hat{\beta}_G$. $n^{3/2}(\hat{\gamma}_S\hat{\alpha}_M\hat{\beta}_G - \gamma_S\alpha_M\beta_G)\sigma_{UVW}^{-1}$ converges to the standard normal distribution where $\sigma_{UVW}^2 = n^2\gamma_S^2\alpha_M^2\sigma_\beta^2 + n^2\gamma_S^2\beta_G^2\sigma_\alpha^2 + n^2\alpha_M^2\beta_G^2\sigma_\gamma^2$ ($\to \infty$ as $n \to \infty$) and construct test statistic $T_{UVW} = \frac{(\hat{\gamma}_S\hat{\alpha}_M\hat{\beta}_G - 0)^2}{\gamma_S^2\alpha_M^2\sigma_{\beta n}^2 + \gamma_S^2\beta_G^2\sigma_{\alpha n}^2 + \alpha_M^2\beta_G^2\sigma_{\gamma n}^2}$, where $n\sigma_{\gamma n}^2 \to \sigma_\gamma^2$, $n\sigma_{\alpha n}^2 \to \sigma_\alpha^2$, $n\sigma_{\beta n}^2 \to \sigma_\beta^2$. T_{UVW} asymptotically follows the central χ^2 distribution with 1 degree of freedom (DF) under the null (PT-N*). By plugging in $\hat{\gamma}_S$, $\hat{\alpha}_M$ and $\hat{\beta}_G$, we obtain $\hat{T}_{UVW} = \frac{(\hat{\gamma}_S\hat{\alpha}_M\hat{\beta}_G - 0)^2}{\hat{\gamma}_S^2\hat{\alpha}_M^2\sigma_{\beta n}^2 + \hat{\gamma}_S^2\hat{\beta}_G^2\sigma_{\alpha n}^2 + \hat{\alpha}_M^2\hat{\beta}_G^2\sigma_{\gamma n}^2}$ (PT-N).

We next propose a joint significance test to test null (7).

DEFINITION 2 (JT of $U_n V_n W_n$). Conduct two-sided z-tests on $\frac{\sqrt{n}|\hat{\gamma}_S|}{\sqrt{n\sigma_{\gamma n}^2}}$, $\frac{\sqrt{n}|\hat{\alpha}_M|}{\sqrt{n\sigma_{\alpha n}^2}}$, and $\frac{\sqrt{n}|\hat{\beta}_G|}{\sqrt{n\sigma_{\beta n}^2}}$ by comparing with $N(0,1)$, and obtain the largest p-value from the three tests as the p-value of the joint significance test of null (7).

THEOREM 3.2. *If $\frac{\gamma_S^2\alpha_M^2\sigma_{\beta n}^2 + \gamma_S^2\beta_G^2\sigma_{\alpha n}^2 + \alpha_M^2\beta_G^2\sigma_{\gamma n}^2}{\hat{\gamma}_S^2\hat{\alpha}_M^2\sigma_{\beta n}^2 + \hat{\gamma}_S^2\hat{\beta}_G^2\sigma_{\alpha n}^2 + \hat{\alpha}_M^2\hat{\beta}_G^2\sigma_{\gamma n}^2} \to 1$ as $n \to \infty$, and at least two of γ_S, α_M, and β_G are nonzero, the joint significance test of $\hat{\gamma}_S$, $\hat{\alpha}_M$, and $\hat{\beta}_G$ in Definition 2 is asymptotically more powerful than the product significance test based on $\hat{T}_{UVW}$.*

Causal assumptions for identifying $\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_M$ and $\gamma_S\alpha_M\beta_G$ are discussed in the Supplementary Material[Huang (2018)].

4. Size of joint significance test. In this section, we show that the proposed JTs have proper test sizes. We first establish that JT is an Intersection-Union Test (IUT), a test with a rejection region of the form $R = \bigcap_{k=1}^K R_k$, where R_k is the rejection region for a test of $H_{0k} : \theta \in \Theta_k$ [Berger and Hsu (1996)]. For JT of $\mathbf{U}_n^T\mathbf{V}_n$ in Definition 1 for the null (3): $H_0 : \bigcup_{k=1,\dots,2^p}\Theta_{\mathbf{U}_n^T\mathbf{V}_n, k}$ is obviously an IUT because $\{\frac{(\hat{\boldsymbol{\alpha}}_S^T\hat{\boldsymbol{\beta}}_M)^2}{V_k} > \kappa\}$ where V_k is an element of $\{\mathbf{w}^{*T}\Sigma_n\mathbf{w}^*\}$ corresponds to the rejection region R_k for $\Theta_{\mathbf{U}^T\mathbf{V}, k}$ and κ is a cut-off value. For JT of $U_n V_n W_n$ in Definition 2, null (7) is equivalent to

$$H_0 : \Theta_{UVW1} \cup \Theta_{UVW2} \cup \Theta_{UVW3},$$

where $\Theta_{UVW1} = \{\theta = (\gamma_S, \alpha_M, \beta_G)^T : \gamma_S = 0, \alpha_M \in \Re, \beta_G \in \Re\}$, $\Theta_{UVW2} = \{\theta : \gamma_S \in \Re, \alpha_M = 0, \beta_G \in \Re\}$, $\Theta_{UVW3} = \{\theta : \gamma_S \in \Re, \alpha_M \in \Re, \beta_G = 0\}$, and the rejection region is $R_{UVWV1} \cap R_{UVW2} \cap R_{UVW3}$ where R_{UVW1}, R_{UVW2}, and R_{UVW3} are the rejection regions for tests of Θ_{UVW1}, Θ_{UVW2}, and Θ_{UVW3}, respectively. Therefore, under the corresponding null, the two proposed JTs are IUTs, which have been shown level α tests, that is, $\sup_{\theta \in \Theta_0} \pi(\theta) \leq \alpha$ [Berger and Hsu (1996)], where $\pi(\theta)$ is the power function and $\Theta_0 = \bigcup_{k=1,\dots,2^p} \Theta_{\mathbf{U}_n^T \mathbf{V}_n, k}$ or $\bigcup_{k=1,2,3} \Theta_{UVWk}$. We show in the Appendix that they are also size α tests.

THEOREM 4.1.   *Under their corresponding null hypotheses, joint significance tests in Definitions* 1 *and* 2 *are intersection-union tests and size α tests, that is,* $\sup_{\theta \in \Theta_0} \pi(\theta) = \alpha$.

In product significance test, the strong effect of a parameter amplifies the effect of the other parameter since, for example, $\hat{\alpha}_S \hat{\beta}_M$ are examined as a product in the numerator of test statistic $\hat{T}_{\mathbf{U}^T \mathbf{V}}$. However, it pays the price by introducing more variability in the denominator of $\hat{T}_{\mathbf{U}^T \mathbf{V}}$ and ends up with lower power than joint significance test, as shown in Theorem 3.1. The separate testing in JT serves as a mechanism to protect Type I Error rate. For example, if $\alpha_S = 0$ and $\beta_M \gg 0$, a test that tends to reject the null by being affected by the very large $\hat{\beta}_M$ may inflate Type I Error rate. In this case, JT is dominated by testing $\alpha_S = 0$, and thus is not likely to reject the null and has well-protected Type I Error rate.

5. Simulation.   Numerical experiment is conducted to demonstrate Lemmas 3.1 and 3.2. For the univariate version of Lemma 3.1, U_n and V_n are generated repeatedly for 10,000 times, from two independent Gaussian distributions with respective means $\sqrt{n} \times 0.5$ and $\sqrt{n} \times 0.2$ and standard deviations (SD) 1 and 2. By increasing n, we show that $U_n V_n - n \times 0.5 \times 0.2$ standardized by its limiting variance $n \times 0.5^2 \times 2^2 + n \times 0.2^2 \times 1^2$ converges to standard normal (Figure 2a). For the multivariate setting, $\mathbf{U}_n$ and $\mathbf{V}_n$ are generated repeatedly for 10,000 times, from two independent multivariate Gaussian distributions with respective means $\sqrt{n} \times (0.5, 0.25, 0.25, 0.75, 0.75)^T$ and $\sqrt{n} \times (0.2, 0.1, 0.1, 0.3, 0.3)^T$ and covariance matrices Σ_U and Σ_V where Σ_U is a 5-by-5 matrix with 1 on the diagonal and 0.3 on the off-diagonal and Σ_V is another 5-by-5 matrix with 1 on the diagonal and 0.5 on the off-diagonal. By increasing n, we observe that $\mathbf{U}_n^T \mathbf{V}_n$ standardized by mean and variance converges to standard normal (Figure 2b). For Lemma 3.2, 10,000 U_n, V_n, and W_n are generated from three independent Gaussian distributions with respective means $\sqrt{n} \times 0.5$, $\sqrt{n} \times 0.2$, and $\sqrt{n} \times 0.3$ and SD 1, 2, and 0.5. We show in Figure 2c that standardized $U_n V_n W_n$ converges to standard normal.
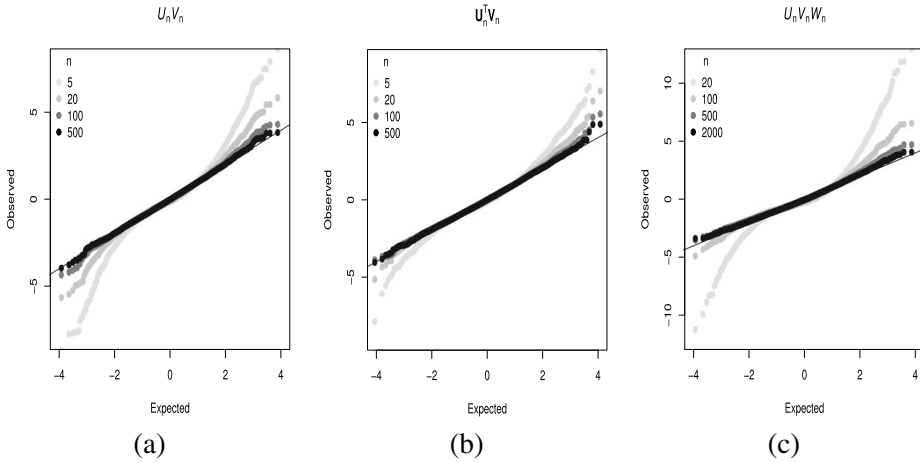
FIG. 2. *Quantile–quantile (QQ) plot of $U_n V_n$, $\mathbf{U}_n^T \mathbf{V}_n$, and $U_n V_n W_n$ standardized by mean and variance against standard normal. U_n, V_n, W_n, $\mathbf{U}_n$, and $\mathbf{V}_n$ follow Gaussian distributions described in Section 5, and n is the scaling parameter.*

5.1. *Test of $\mathbf{U}_n^T \mathbf{V}_n$.* To evaluate the finite-sample performance of tests for null hypothesis (3), we generate S from standard normal and Y and $\mathbf{M}$ according to the models: $Y_i = 1.2 + 0.5 \times S_i + \sum_j M_{ji} \beta_{Mj} + \varepsilon_{Yi}$ and $M_{ji} = 0.8 + S_i \times \alpha_{Sj} + \varepsilon_{Mji}$, where $j = 1, \ldots, 5$, $i = 1, \ldots, n$, $\alpha_{Sj} = 0.8 \times \delta$, $\beta_{Mj} = 0.4 \times \delta$, $\boldsymbol{\alpha}_S = (\alpha_{S1}, \ldots, \alpha_{S5})^T$, $\boldsymbol{\beta}_M = (\beta_{M1}, \ldots, \beta_{M5})^T$; ε_{Yi} follows standard normal; $\boldsymbol{\varepsilon}_{Mi} = (\varepsilon_{M1i}, \ldots, \varepsilon_{M5i})^T$ follows a multivariate normal with zero mean and a 5-by-5 covariance matrix with 0.8 on the diagonal and 0 on the off-diagonal; and ε_{Yi} and $\boldsymbol{\varepsilon}_{Mi}$ are independent. For the null, $\alpha_{S4} = \alpha_{S5} = \beta_{M1} = \beta_{M2} = \beta_{M3} = 0$, $\alpha_{S1} = \alpha_{S2} = \alpha_{S3} = 0.48$, and $\beta_{M4} = \beta_{M5} = 0.24$ (Figure 3). Two sets of simulation are conducted under the alternative: one with $n = 2000$ and δ increasing from 0.005 to 0.4 (Figures 4a and b), and the other with $\delta = 0.4$ and n from 20 to 500 (Figures 4c and d). Three versions of product significance test are implemented: PT-N approximates the mediation effect using Gaussian distributions as specified in Remark 1 (and 2); and PT-NP approximates the normal product distribution with $\{\alpha^{(b)} \beta^{(b)} - \hat{\alpha}_S \hat{\beta}_M\}$ where $\alpha^{(b)}$ and $\beta^{(b)}$ were resampled repeatedly from $N(\hat{\alpha}_S, \hat{\sigma}_{\alpha n}^2)$ and $N(\hat{\beta}_M, \hat{\sigma}_{\beta n}^2)$, respectively [Huang and Pan (2016)].

Under the null, the test statistic of JT has a faster convergence to χ_1^2 distribution with the same configuration as the sample size increases, compared to PT-N (Figure 3). With $n = 10$, $p = 5$, and $\delta = 0.4$, JT protects Type I Error rate, and distribution of the test statistic $\frac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{V_{\max}}$ approaches χ_1^2 distribution (Figures 3c and d). In contrast, PT-N requires $n \geq 100$ to ensure a descent convergence (Figure 3b). Under the null where we set $\beta_{M1} = \beta_{M2} = \beta_{M3} = \alpha_{S4} = \alpha_{S5} = 0$ and the rest of nonzero parameters as $\alpha_{Sj} = 0.8 \times \delta$ and $\beta_{Mj} = 0.4 \times \delta$, Type I Error rate of all four tests approaches 5%, and standardized $\hat{T}_{\mathbf{U}^T \mathbf{V}}$ is close to normal when $\delta \geq 0.12$
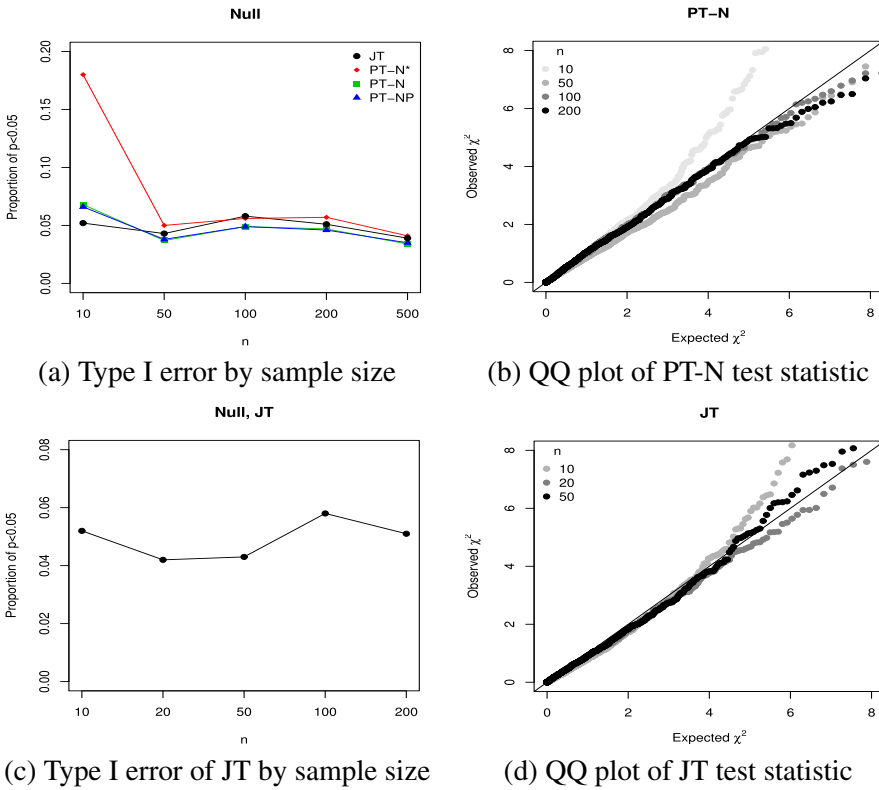
(a) Type I error by sample size

(b) QQ plot of PT-N test statistic

(c) Type I error of JT by sample size

(d) QQ plot of JT test statistic

FIG. 3.   *Simulation results of tests for* $\mathbf{U}_n^T \mathbf{V}_n$ *under the null. Proportion of* $p < 0.05$ *by sample size n of mediation tests is depicted for the four tests* (a) *and for JT with a narrower range* (c). *Test statistics of PT-N,* $\dfrac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{\hat{\boldsymbol{\alpha}}_S^T \Sigma_{\boldsymbol{\beta}n} \hat{\boldsymbol{\alpha}}_S + \hat{\boldsymbol{\beta}}_M^T \Sigma_{\boldsymbol{\alpha}n} \hat{\boldsymbol{\beta}}_M}$ *and JT,* $\dfrac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{V_{\max}}$ *are evaluated against* χ^2 *distribution with 1 DF* [(b) *and* (d)].

[see the Supplementary Material Huang (2018)]. Under the null of small nonzero α_{Sj} and β_{Mj}, $\hat{T}_{\mathbf{U}^T\mathbf{V}}$ (PT-N) does not follow normal distribution, which explains the very conservative test size of PT-N.

Due to the poor normal approximation of PT-N with small effect or sample size, the cut-off for the power simulation is adjusted to ensure 5% error rate. Under the alternative, standardized $n\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M$ is close to standard normal as $\delta \geq 0.052$ or $n \geq 50$ (Figures 4b and d). The proposed JT has the highest power when the normal approximation for PT gets better, that is, $\delta \geq 0.052$ (Figure 4a) or $n \geq 50$ (Figure 4c). PT-NP has similar performance to PT-N. Simulation in the univariate setting, $U_n V_n$ is presented in the Supplementary Material [Huang (2018)].

5.2. *Test of* $U_n V_n W_n$. We simulate S from standard normal, and M, G, and Y with the following models: $M_i = 0.8 + S_i \times \gamma_S + \varepsilon_{Mi}$, $G_i = 1.2 + 0.5 \times S_i +$
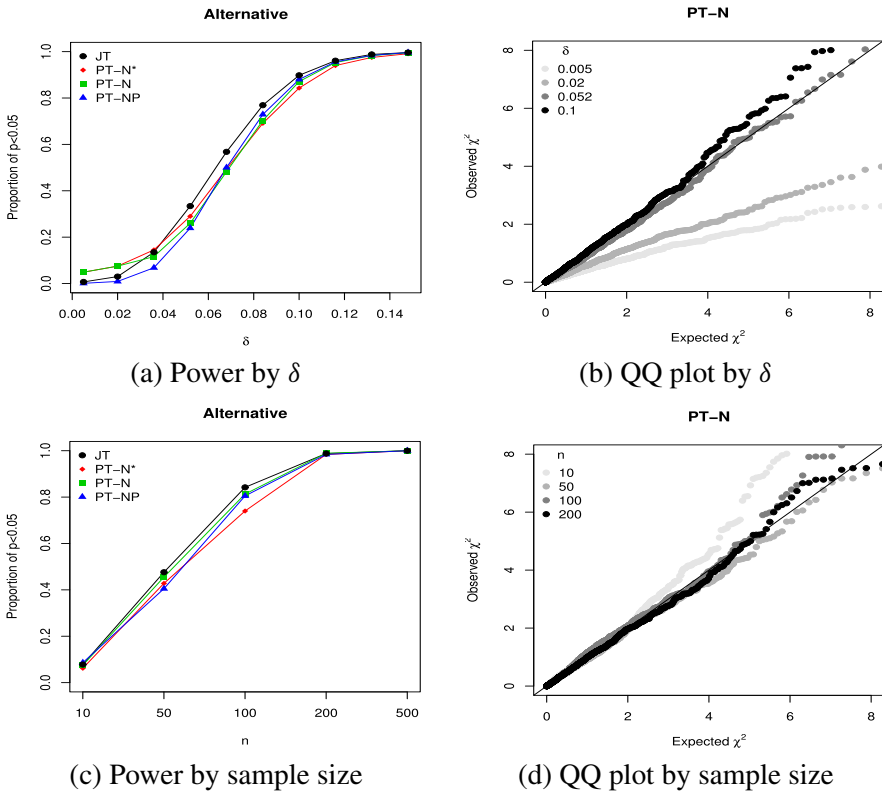
FIG. 4.   *Simulation results of tests for $\mathbf{U}_n^T \mathbf{V}_n$ under the alternative. Proportion of $p < 0.05$ by effect size indexed by δ (a) and by sample size n (c) of mediation tests is depicted. The distribution of $\hat{T}_{\mathbf{U}^T \mathbf{V}}$ is evaluated against χ^2 distribution with 1 DF [(b) and (d)].*

$M_i \times \alpha_M + \varepsilon_{Gi}$ and $Y_i = 1.0 + 0.5 \times S_i + 0.3 \times M_i + G_i \times \beta_G + \varepsilon_{Yi}$, where $\gamma_S = 0.8 \times \delta$, $\alpha_M = 0.6 \times \delta$, $\beta_G = 0.5 \times \delta$, and ε_{Mi}, ε_{Gi} and ε_{Yi} are independent normal random variables with zero mean and SD 0.8. For the null, we study two scenarios: one with varying n, and $\alpha_M = 0$, $\gamma_S = 0.48$, $\beta_G = 0.3$ (Figure 5), and the other with varying δ, and $\alpha_M = 0$, $\gamma_S = 0.8 \times \delta$, $\beta_G = 0.5 \times \delta$, $n = 1000$ [see the Supplementary Material Huang (2018)]. We conduct two sets of simulation under the alternative: one with $n = 2000$ and $\delta = 0.005$ to 0.8 (Figures 6a and b), and the other with $\delta = 0.4$ and $n = 50$ to 1000 (Figures 6c and d).

Under the null, the test statistic of JT converges to χ_1^2 distribution faster than that of PT-N as n increases (Figure 5). PT-N and JT protect Type I Error rate with $n \geq 1000$ and $n \geq 200$, respectively (Figures 5a and c), and their test statistics approach χ_1^2 distribution with the similar sample sizes (Figures 5b and d). The cut-offs for the power simulation of normality-based tests are adjusted by the test size due to the poor approximation under small effect or sample size. Under the alternative, we observe that standardized $n^{3/2} \hat{\gamma}_S \hat{\alpha}_M \hat{\beta}_G$ approaches normality as n or δ increases.
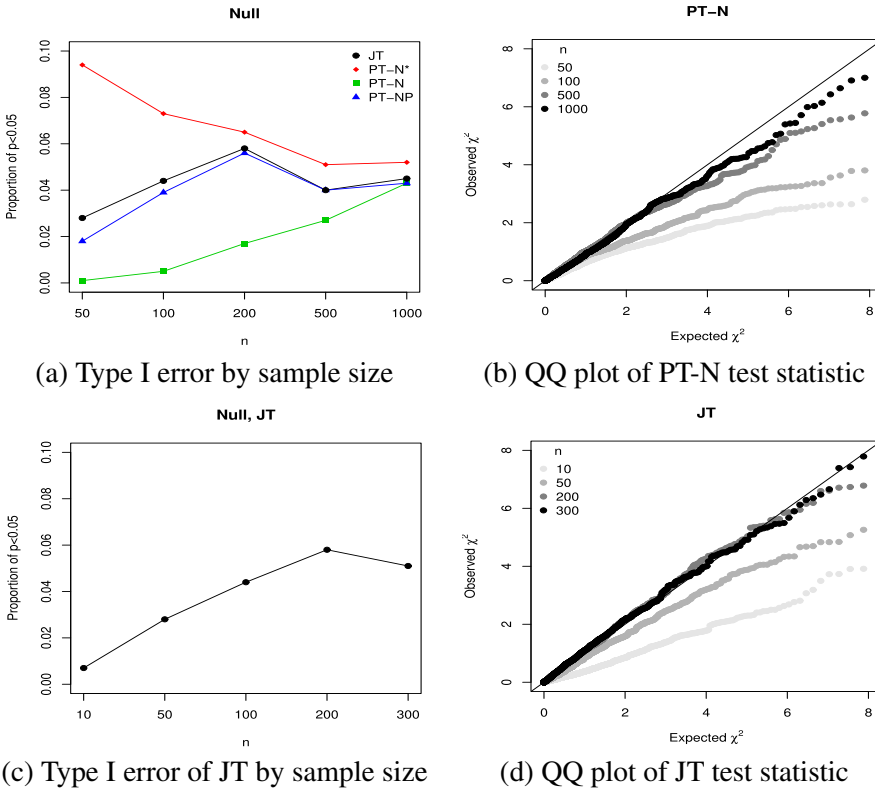
(a) Type I error by sample size

(b) QQ plot of PT-N test statistic

(c) Type I error of JT by sample size

(d) QQ plot of JT test statistic

FIG. 5. *Simulation results of tests for $U_n V_n W_n$ under the null. Proportion of $p < 0.05$ by sample size n of mediation tests is depicted for the four tests* (a) *and for JT with a narrower range* (c). *Test statistics of PT-N*, $\frac{(\hat{\gamma}_S \hat{\alpha}_M \hat{\beta}_G)^2}{\hat{\gamma}_S^2 \hat{\alpha}_M^2 \sigma_{\beta n}^2 + \hat{\gamma}_S^2 \hat{\beta}_G^2 \sigma_{\alpha n}^2 + \hat{\alpha}_M^2 \hat{\beta}_G^2 \sigma_{\gamma n}^2}$ *and JT*, $\min\{\frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}, \frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}, \frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}\}$ *are evaluated against χ^2 distribution with* 1 *DF* [(b) *and* (d)].

JT has the highest power among the four tests as normal approximation for PT improves, and PT-NP and JT have very similar performance (Figure 6). Under small n (≤ 100) or δ (≤ 0.24), $\hat{T}_{UVW}$ has a severe departure from normality, which leads to the conservative Type I Error rate [see the Supplementary Material Huang (2018)].

In summary, the numerical experiment supports the theoretical results in Sections 3 and 4. Normal approximation of PT works better for larger effects or sample size. Under the finite sample, joint significance tests have better power than normality-based PT, and the performance of NP-based product significance test is in between.

6. Mediation analysis of NEFS. DNA methylation data are preprocessed using `methylumi` package in R [Davis et al. (2015)] and normalized using Beta-Mixture Quantile Dilation approach [Teschendorff et al. (2013)]. The processed
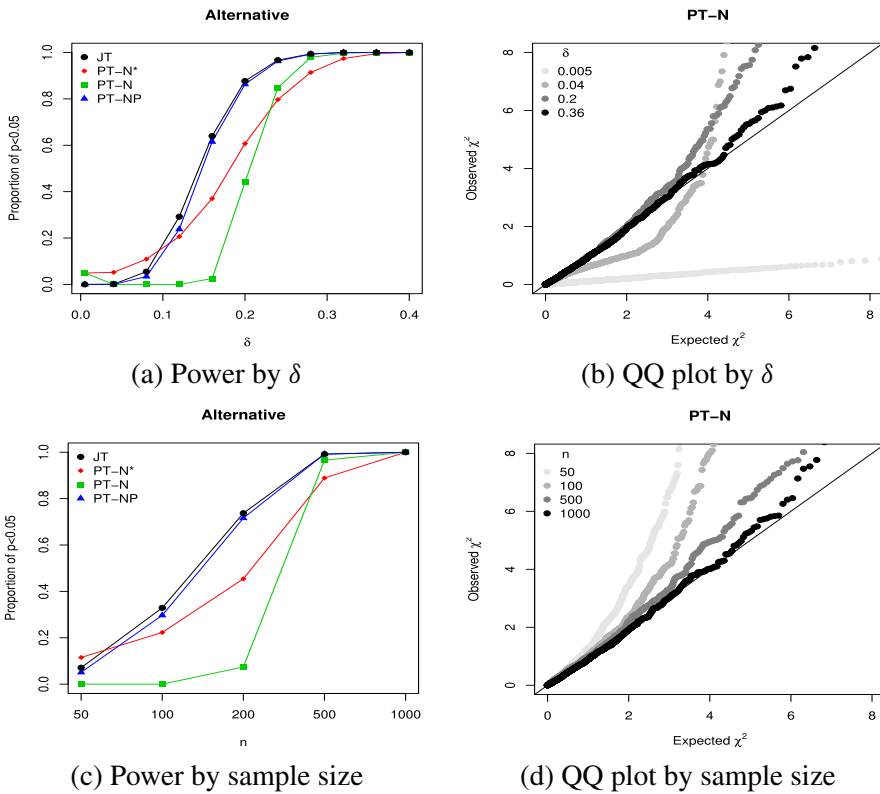
FIG. 6. *Simulation results of tests for $U_n V_n W_n$ under the alternative. Proportion of $p < 0.05$ by effect size of the nonzero parameter indexed by δ (a) and by sample size n (c) of mediation tests is depicted. The distribution of $\hat{T}_{UVW}$ is evaluated by QQ plots against χ^2 distribution with 1 degree of freedom by δ [(b)] and sample size [(d)].*

methylation values ranging from 0 to 1 are then logit transformed prior to analyses.

We first set up a mediation model in Figure 1a where S is the socioeconomic index at age 7, M is methylation level of a locus in chromosome 17, and Y is BMI at adulthood. We chose to focus on chromosome 17 because (1) it harbors an interesting gene *FASN* (fatty acid synthase) that has been previously reported for its association with BMI, and (2) the computation cost for analyzing all chromosomes was too high, especially for the bootstrap-based method. JT and three PTs were conducted to analyze 16,394 methylation loci in chromosome 17 one at a time; and both PT-NP and PT-B were carried out with 1000 replicates; PT-B approximates the distribution of $\hat{\alpha}_S \hat{\beta}_M$ through bootstrapping: $\{\hat{\alpha}_S^{(b)} \hat{\beta}_M^{(b)} - \frac{1}{B} \sum_{b=1}^{B} \hat{\alpha}_S^{(b)} \hat{\beta}_M^{(b)}\}$, where $\hat{\alpha}_S^{(b)}$ and $\hat{\beta}_M^{(b)}$ are estimated from the bootstrap data, sampled from the original data with replacement, and B is the number of bootstrapping, 1000. There are
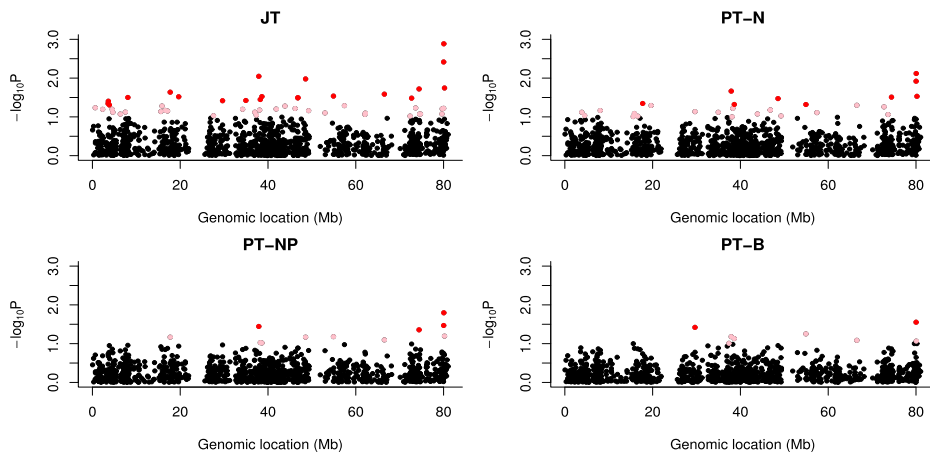
FIG. 7. *Gene-based mediation analysis of* 1151 *genes on chromosome* 17 *using different tests. Red*: $p < 0.05$; *pink*: $0.05 \leq p < 0.1$.

243, 24, 193, and 177 methylation loci with $p < 0.05$ for JT, PT-N, PT-NP, and PT-B, respectively; computation time is shown in Table 1.

We were interested to examine whether all methylation loci within a gene *en bloc* mediate the effect of socioeconomic disadvantage on BMI (Figure 1b) utilizing the multi-mediator analyses in Section 3.1. We conducted the gene-based analysis, examining 1151 genes with multiple methylation level measures (< 30 loci) one at a time (Figure 7). To satisfy the condition $\alpha_{Sj}\alpha_{Sk}\rho_{\beta,jk} = 0$ and $\beta_{Mj}\beta_{Mk}\rho_{\alpha,jk} = 0$, we transformed the methylation levels to be uncorrelated by carrying out singular value decomposition on the sample covariance of the residuals of model (2) and performed the proposed tests on the transformed mediators. Using JT, PT-N, PT-NP, and PT-B, respectively, 21, 9, 4, and 2 genes have p-values < 0.05, and 51, 28, 11, and 8 genes have $p < 0.1$. *FASN* is the most significant gene with 11, 3, 11, and 10 loci having single-locus $p < 0.05$ using JT, PT-N, PT-NP, and PT-B, respectively (Figure 8); respective p-values of the gene-based overall mediation effect are 0.0013, 0.0076, 0.028, and 0.086. PT-B took 1.58 hours for the 1151 gene-based analyses, compared to 7.52 seconds for JT (Table 1).

We investigated whether the significant mediation effect by *FASN* is further mediated by the childhood BMI. None of the 26 loci revealed such a path-specific effect (see Supplementary Material Table S1), and cg04029737 had the smallest p-value: 0.297 in JT, 0.322 in PT-N, 0.318 in PT-NP, and 0.390 in PT-B. JT had the smallest p-value among the four tests in 19 out of 26 loci. The results suggest that DNA methylation level of *FASN* gene mediates the effect of socioeconomic disadvantage on BMI in women, which may not be mediated by the childhood BMI.
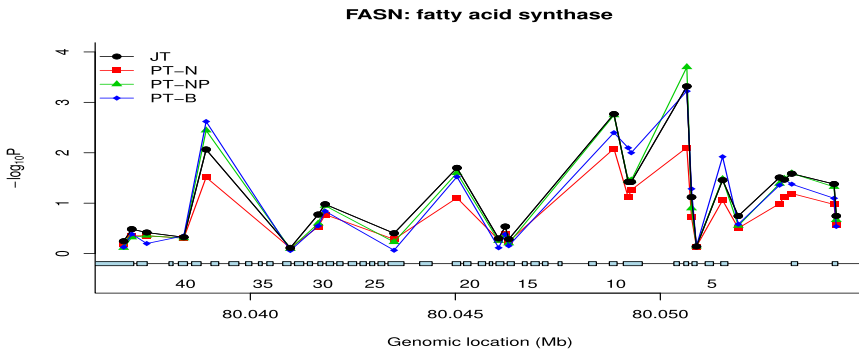
FIG. 8.   *p-values for mediation effects of* 26 *methylation loci within FASN gene. The boxes at the bottom indicate exons of the gene and the numerals indicate the exon numbers.*

Studies have shown that *FASN* is involved in regulation of body weight [Kovacs et al. (2004), Loftus et al. (2000)]. In mice, *FASN* expression in adipose tissue can be affected by high dietary fat [Kadota et al. (2016)]. Compared to those with high socioeconomic status, individuals with lower socioeconomic status in industrialized nations consume less methyl donor foods such as fruits and vegetables [Darmon and Drewnowski (2008), Giskes et al. (2010)]. Our analyses support the mediation mechanism that socioeconomic adversity may affect the dietary habit to alter epigenetics of *FASN*, and through regulating gene expression, *FASN* epigenetics can further determine body weight.

We conclude that joint significance test is a powerful and very computationally efficient test (Table 1). Implementation codes are available at http://www.stat.sinica.edu.tw/ythuang/MMT-JT.zip.

7. Discussion.   Test size is an upper limit across all parameters in the null space. As illustrated in simulation studies, the joint significance tests are conservative if the effect is small and the target size α is achieved if the nonzero parameter is large. Such a property affects its utility in high-dimensional data analyses. When applying multiplicity adjustment procedures against theoretical distributions with exponential or geometric decay of effect size or some empirical distributions close

TABLE 1
Computation time (*in seconds unless specified*) *with* 3.40 *GHz CPU and* 16.0 *GB RAM.* 1000
replicates for PT-NP and PT-B

	No. of analyses	JT	PT-N	PT-NP	PT-B
Single-locus analyses (UV)	16,394	64.4	64.1	67.4	16.0 hours
Gene-based analyses ($\mathbf{U}^T\mathbf{V}$)	1151	7.52	7.46	11.84	1.58 hours
Path-specific effect (UVW)	26	0.08	0.13	0.09	142

to it, the correction would be very conservative. How to account for multiple testing issue when the majority of signals for both parameters are zero warrants further studies.

We provide a general insight about why JT is more powerful than PT. In the univariate mediation effect of $\alpha_S \beta_M$, PT with normal approximation tests $\sqrt{n}\hat{\alpha}_S\sqrt{n}\hat{\beta}_M$ where both $\sqrt{n}\hat{\alpha}_S$ and $\sqrt{n}\hat{\beta}_M$ vary in their respective distributions; JT examines $\hat{\alpha}_S\sqrt{n}\hat{\beta}_M$ ($\to \alpha_S\sqrt{n}\hat{\beta}_M$) and $\hat{\beta}_M\sqrt{n}\hat{\alpha}_S$ ($\to \beta_M\sqrt{n}\hat{\alpha}_S$), and both converges to a normal distribution multiplied by a constant. The larger variability in PT can be explicitly expressed by the inequality in the Appendix, which guarantees that either test of $\hat{\alpha}_S\sqrt{n}\hat{\beta}_M$ and $\hat{\beta}_M\sqrt{n}\hat{\alpha}_S$ is more powerful than the PT, and the validity of JT in combining the separate tests by picking the least significant one is supported by the theory of IUT.

APPENDIX

PROOF OF LEMMA 3.1.   The moment generating function (mgf) of $\mathbf{U}_n^T\mathbf{V}_n$ is

$M_{\mathbf{U}_n^T\mathbf{V}_n}(t)$

$$
= \frac{1}{2\pi|\Sigma_1|^{1/2}|\Sigma_2|^{1/2}}
$$
$$
\times \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\mathbf{u}^T\mathbf{v}t} e^{-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu}_{1n})^T\Sigma_1^{-1}(\mathbf{u}-\boldsymbol{\mu}_{1n})-\frac{1}{2}(\mathbf{v}-\boldsymbol{\mu}_{2n})^T\Sigma_2^{-1}(\mathbf{v}-\boldsymbol{\mu}_{2n})}\, d\mathbf{u}\, d\mathbf{v}
$$
$$
= |\mathbf{I}-t^2\Sigma_1\Sigma_2|^{-1/2}
$$
$$
\times e^{\frac{1}{2}t^2\boldsymbol{\mu}_{1n}^T(\mathbf{I}-t^2\Sigma_1\Sigma_2)^{-1}\Sigma_2\boldsymbol{\mu}_{1n}+\frac{1}{2}t^2\boldsymbol{\mu}_{2n}^T(\mathbf{I}-t^2\Sigma_1\Sigma_2)^{-1}\Sigma_1\boldsymbol{\mu}_{2n}+t\boldsymbol{\mu}_{1n}^T(\mathbf{I}-t^2\Sigma_1\Sigma_2)^{-1}\boldsymbol{\mu}_{2n}}.
$$

With $\boldsymbol{\mu}_{1n}^T\Sigma_2\boldsymbol{\mu}_{1n} + \boldsymbol{\mu}_{2n}^T\Sigma_1\boldsymbol{\mu}_{2n} \to \infty$, it can be easily shown that the mgf of $Z \equiv \frac{\mathbf{U}_n^T\mathbf{V}_n - \boldsymbol{\mu}_{1n}^T\boldsymbol{\mu}_{2n}}{\sqrt{\boldsymbol{\mu}_{1n}^T\Sigma_2\boldsymbol{\mu}_{1n}+\boldsymbol{\mu}_{2n}^T\Sigma_1\boldsymbol{\mu}_{2n}}}$, $M_Z(t) \to e^{\frac{1}{2}t^2}$, the mgf of standard normal distribution. Due to the uniqueness of mgf, Z converges to standard normal distribution.   □

PROOF OF THEOREM 3.1.   Because $\frac{\boldsymbol{\alpha}_S^T\Sigma_{\beta n}\boldsymbol{\alpha}_S+\boldsymbol{\beta}_M^T\Sigma_{\alpha n}\boldsymbol{\beta}_M}{\hat{\boldsymbol{\alpha}}_S^T\Sigma_{\beta n}\hat{\boldsymbol{\alpha}}_S+\hat{\boldsymbol{\beta}}_M^T\Sigma_{\alpha n}\hat{\boldsymbol{\beta}}_M} \to 1$ in probability, under the null, $\hat{T}_{\mathbf{U}^T\mathbf{V}} = T_{\mathbf{U}^T\mathbf{V}}\frac{\boldsymbol{\alpha}_S^T\Sigma_{\beta n}\boldsymbol{\alpha}_S+\boldsymbol{\beta}_M^T\Sigma_{\alpha n}\boldsymbol{\beta}_M}{\hat{\boldsymbol{\alpha}}_S^T\Sigma_{\beta n}\hat{\boldsymbol{\alpha}}_S+\hat{\boldsymbol{\beta}}_M^T\Sigma_{\alpha n}\hat{\boldsymbol{\beta}}_M}$ converges to central χ^2 distribution with 1 DF. Under the assumption $\alpha_{Sj}\alpha_{Sk}\rho_{\beta,jk} \geq 0$ and $\beta_{Mj}\beta_{Mk}\rho_{\alpha,jk} \geq 0$, each term in $\hat{\boldsymbol{\alpha}}_S^T n\Sigma_{\beta n}\hat{\boldsymbol{\alpha}}_S + \hat{\boldsymbol{\beta}}_M^T n\Sigma_{\alpha n}\hat{\boldsymbol{\beta}}_M = n\sum_j \hat{\alpha}_{Sj}^2\sigma_{\beta nj}^2 + n\sum_{j\neq k}\hat{\alpha}_{Sj}\hat{\alpha}_{Sk}\rho_{\beta njk} + n\sum_j \hat{\beta}_{Mj}^2\sigma_{\alpha nj}^2 + n\sum_{j\neq k}\hat{\beta}_{Mj}\hat{\beta}_{Mk}\rho_{\alpha njk}$ converges to a nonnegative value, and $nV_{\max}$ is a subset of the above $2p^2$ terms. Therefore, asymptotically

$$
\hat{T}_{\mathbf{U}^T\mathbf{V}} = \frac{(\sqrt{n}\hat{\boldsymbol{\alpha}}_S^T\hat{\boldsymbol{\beta}}_M)^2}{n\hat{\boldsymbol{\alpha}}_S^T\Sigma_{\beta n}\hat{\boldsymbol{\alpha}}_S + n\hat{\boldsymbol{\beta}}_M^T\Sigma_{\alpha n}\hat{\boldsymbol{\beta}}_M} \leq \frac{(\sqrt{n}\hat{\boldsymbol{\alpha}}_S^T\hat{\boldsymbol{\beta}}_M)^2}{nV_{\max}}.
$$

Note the two sides of the above inequality are the test statistics of PT-N and JT. If the two statistics follow the same asymptotic distribution under the null, the inequality implies that the proposed JT always has a smaller p-value than the test based on $\hat{T}_{\mathbf{U}^T\mathbf{V}}$ (PT-N), which guarantees that JT is statistically more powerful than PT-N: $\pi_{PT-N}(\theta) = P(\hat{T}_{\mathbf{U}^T\mathbf{V}} > \kappa) < P(\frac{(\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{V_{\max}} > \kappa) = \pi_{JT}(\theta)$ due to the above inequality, where $\Phi(\sqrt{\kappa}) = 1 - \alpha/2$ and $\Phi(\cdot)$ is the cumulative distribution function of standard normal. We show in the following that the test statistic of JT asymptotically also follows χ^2 distribution with 1 DF under the null.

Denote $\boldsymbol{\alpha}_S = (\boldsymbol{\alpha}_{S1}^T, \boldsymbol{\alpha}_{S2}^T)^T$ and $\boldsymbol{\beta}_M = (\boldsymbol{\beta}_{M1}^T, \boldsymbol{\beta}_{M2}^T)^T$, where $\boldsymbol{\alpha}_{S1} = (\alpha_{S1}, \ldots, \alpha_{Sp/2})^T$ and $\boldsymbol{\beta}_{M1} = (\beta_{M1}, \ldots, \beta_{Mp/2})^T$. Without loss of generality, we focus on the two different types of null hypotheses: (1) $\boldsymbol{\alpha}_{S1} = \mathbf{0}$, $\boldsymbol{\alpha}_{S2} \neq \mathbf{0}$, $\boldsymbol{\beta}_{M1} \neq \mathbf{0}$, $\boldsymbol{\beta}_{M2} = \mathbf{0}$ and (2) $\alpha_{S1} \neq 0$, $(\alpha_{S2}, \ldots, \alpha_{Sp})^T = \mathbf{0}$ and $\boldsymbol{\beta}_M = \mathbf{0}$. Under the null of $\boldsymbol{\alpha}_{S1} = \mathbf{0}$, $\boldsymbol{\alpha}_{S2} \neq \mathbf{0}$, $\boldsymbol{\beta}_{M1} \neq \mathbf{0}$, and $\boldsymbol{\beta}_{M2} = \mathbf{0}$, $nV_{\max} \to \mathbf{w}_0^{*T} \Sigma \mathbf{w}_0^*$, where $\mathbf{w}_0^{*T} = (\boldsymbol{\beta}_{M1}^T, \mathbf{0}^T, \mathbf{0}^T, \boldsymbol{\alpha}_{S2}^T)$ and $n\Sigma_n \to \Sigma = \left[\begin{smallmatrix} \Sigma_\alpha & \mathbf{0} \\ \mathbf{0} & \Sigma_\beta \end{smallmatrix}\right]$, a bounded covariance matrix. Since $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \to N(\mathbf{0}, \Sigma)$, one can show that $\sqrt{n}\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M = \sqrt{n}\mathbf{w}^{*T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \to N(0, \mathbf{w}_0^{*T} \Sigma \mathbf{w}_0^*)$. Therefore, $\frac{(\sqrt{n}\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{nV_{\max}}$ follows χ^2 distribution with 1 DF under the null.

Under the null of $\alpha_{S1} \neq 0$, $(\alpha_{S2}, \ldots, \alpha_{Sp})^T = \mathbf{0}$ and $\boldsymbol{\beta}_M = \mathbf{0}$,

$$
nV_{\max} = n\hat{\alpha}_{S1}^2 \sigma_{\beta n1}^2 + \sum_{j=2}^{p} \hat{\alpha}_{Sj}^2 n\sigma_{\beta nj}^2 + \sum_{j \neq k} \hat{\alpha}_{Sj} \hat{\alpha}_{Sk} n\rho_{\beta njk}
$$

$$
+ \sum_{j} \hat{\beta}_{Mj}^2 n\sigma_{\alpha nj}^2 + \sum_{j \neq k} \hat{\beta}_{Mj} \hat{\beta}_{Mk} n\rho_{\alpha njk}
$$

$$
= n\hat{\alpha}_{S1}^2 \sigma_{\beta n1}^2 + o_p(1).
$$

Further, we have $\sqrt{n}\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M \to N(0, \alpha_{S1}^2 \sigma_{\beta 1}^2)$ in distribution, provided that $\sqrt{n} \times (\hat{\boldsymbol{\beta}}_M - \boldsymbol{\beta}_M) \to N(\mathbf{0}, \Sigma_\beta)$ in distribution and $\hat{\boldsymbol{\alpha}}_S \to (\alpha_{S1}, \mathbf{0}^T)^T$ in probability. The JT test statistic

$$
\frac{(\sqrt{n}\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{nV_{\max}} = \frac{(\sqrt{n}\hat{\boldsymbol{\alpha}}_S^T \hat{\boldsymbol{\beta}}_M)^2}{\alpha_{S1}^2 \sigma_{\beta 1}^2} \left(\frac{\alpha_{S1}^2 \sigma_{\beta 1}^2}{\hat{\alpha}_{S1}^2 n\sigma_{\beta n1}^2}\right) \left(\frac{\hat{\alpha}_{S1}^2 n\sigma_{\beta n1}^2}{nV_{\max}}\right)
$$

converges to central χ^2 distribution with 1 DF because the last two parentheses converge to 1 in probability. Therefore, both test statistics of JT and PT-N asymptotically follow central χ^2 distribution with 1 DF under the null. Note the theorem does not apply to the null of $\boldsymbol{\theta} = \mathbf{0}$ because it violates the assumption $n^2 \boldsymbol{\alpha}_S^T \Sigma_{\beta n} \boldsymbol{\alpha}_S + n^2 \boldsymbol{\beta}_M^T \Sigma_{\alpha n} \boldsymbol{\beta}_M \to \infty$ for normal approximation of PT. □

PROOF OF LEMMA 3.2.   The mgf of $U_n V_n W_n$ is

$$
M_{U_n V_n W_n}(t) = \frac{1}{(2\pi)^{3/2} \sigma_1 \sigma_2 \sigma_3}
$$

$$\times \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{uvwt} e^{-\frac{(u-\mu_{1n})^2}{2\sigma_1^2} - \frac{(v-\mu_{2n})^2}{2\sigma_2^2} - \frac{(w-\mu_{3n})^2}{2\sigma_3^2}} \, du\,dv\,dw$$

$$= \frac{1}{\sqrt{2\pi}\sigma_3} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{1 - \sigma_1^2\sigma_2^2 t^2 w^2}}$$

$$\times e^{\frac{2\mu_{1n}\mu_{2n}wt + (\mu_{1n}^2\sigma_2^2 + \mu_{2n}^2\sigma_1^2)w^2 t^2}{2(1 - \sigma_1^2\sigma_2^2 t^2 w^2)}} e^{-\frac{(w-\mu_{3n})^2}{2\sigma_3^2}} \, dw.$$

Let $Z^* \equiv \frac{U_n V_n W_n}{\sqrt{c}}$ where $c = \mu_{1n}^2\mu_{2n}^2\sigma_3^2 + \mu_{1n}^2\mu_{3n}^2\sigma_2^2 + \mu_{2n}^2\mu_{3n}^2\sigma_1^2 \to \infty$. It follows that

$$M_{Z^*}(t) \to \frac{1}{\sqrt{2\pi}\sigma_3} \int_{-\infty}^{+\infty} e^{\mu_{1n}\mu_{2n}wt/\sqrt{c} + \frac{1}{2}(\mu_{1n}^2\sigma_2^2 + \mu_{2n}^2\sigma_1^2)w^2 t^2/c} e^{-\frac{(w-\mu_{3n})^2}{2\sigma_3^2}} \, dw$$

$$= \frac{1}{\sqrt{1 - (\mu_{1n}^2\sigma_2^2\sigma_3^2 + \mu_{2n}^2\sigma_1^2\sigma_3^2)t^2/c}}$$

$$\times e^{\frac{\mu_{1n}\mu_{2n}\mu_{3n}t/\sqrt{c} + \frac{1}{2}(\mu_{1n}^2\mu_{2n}^2\sigma_3^2 + \mu_{1n}^2\mu_{3n}^2\sigma_2^2 + \mu_{2n}^2\mu_{3n}^2\sigma_1^2)t^2/c}{1 - (\mu_{1n}^2\sigma_2^2\sigma_3^2 + \mu_{2n}^2\sigma_1^2\sigma_3^2)t^2/c}}.$$

Further, by letting $Z = Z^* - \frac{\mu_{1n}\mu_{2n}\mu_{3n}}{\sqrt{c}}$, we show that

$$\lim_{\substack{\text{at least two of} \\ \mu_{1n},\mu_{2n},\mu_{3n} \to \infty}} M_Z(t) = e^{\frac{1}{2}t^2},$$

which is the mgf of the standard normal. $\square$

PROOF OF THEOREM 3.2.   As $\hat{\gamma}_S$, $\hat{\alpha}_M$ and $\hat{\beta}_G$ are MLEs, it can be shown that $\hat{T}_{UVW}$ converges to T_{UVW} following central χ^2 distribution with 1 DF under the null. It can further be shown that

$$\hat{T}_{UVW} = \frac{\hat{\gamma}_S^2\hat{\alpha}_M^2\hat{\beta}_G^2}{\hat{\gamma}_S^2\hat{\alpha}_M^2\sigma_{\beta n}^2 + \hat{\gamma}_S^2\hat{\beta}_G^2\sigma_{\alpha n}^2 + \hat{\alpha}_M^2\hat{\beta}_G^2\sigma_{\gamma n}^2} \le \min\left\{\frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}, \frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}, \frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}\right\}.$$

Equality is attained if any of $\hat{\gamma}_S$, $\hat{\alpha}_M$ and $\hat{\beta}_G$ is zero. If one can show that $\min\{\frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}, \frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}, \frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}\}$ follows central χ^2 distribution with 1 DF under the null (see the following), the above inequality implies that the proposed procedure for joint significance test of $\hat{\gamma}_S$, $\hat{\alpha}_M$, and $\hat{\beta}_G$ always has a smaller p-value, and thus better power than the product significance test based on $\hat{T}_{UVW}$ (PT-N).

Under the null of $\gamma_S = 0, \alpha_M \ne 0$, and $\beta_G \ne 0$, $\Pr(\frac{(\sqrt{n}\hat{\gamma}_S)^2}{n\sigma_{\gamma n}^2} < \frac{(\sqrt{n}\hat{\alpha}_M)^2}{n\sigma_{\alpha n}^2}, \frac{(\sqrt{n}\hat{\gamma}_S)^2}{n\sigma_{\gamma n}^2} < \frac{(\sqrt{n}\hat{\beta}_G)^2}{n\sigma_{\beta n}^2}) \to 1$ because $\sqrt{n}\hat{\gamma}_S = O_p(1)$ and both $\sqrt{n}\hat{\alpha}_M\ [= O_p(\sqrt{n})]$ and $\sqrt{n}\hat{\beta}_G$

$[= O_p(\sqrt{n})]$ go to infinity. It follows that $\min\{\frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}, \frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}, \frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}\} = \frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}$ with probability converging to 1. Similarly, under the null of $\gamma_S \neq 0$, $\alpha_M = 0$, and $\beta_G \neq 0$, $\min\{\frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}, \frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}, \frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}\}$ is dominated by $\frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}$; and under the null of $\gamma_S \neq 0$, $\alpha_M \neq 0$, and $\beta_G = 0$, $\min\{\frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}, \frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}, \frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}\}$ is dominated by $\frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}$. $\frac{\hat{\gamma}_S^2}{\sigma_{\gamma n}^2}$, $\frac{\hat{\alpha}_M^2}{\sigma_{\alpha n}^2}$, and $\frac{\hat{\beta}_G^2}{\sigma_{\beta n}^2}$ all follow central χ^2 distribution with 1 DF under their respective null. $\square$

PROOF OF THEOREM 4.1. We apply Theorem 2 in Berger and Hsu (1996) to prove the JT in Definitions 1 and 2 are size α tests. We have shown in Section 4 that the JTs are IUT for $H_0 : \theta \in \Theta_0 \equiv \bigcup_{k=1}^{K} \Theta_k$. To apply the theorem, one needs to show that there exists a sequence of parameters points $\theta_l \in \Theta_i$ ($l = 1, \ldots, \infty$) for some $i = 1, \ldots, K$ such that (i) $\lim_{l \to \infty} P_{\theta_l}(\mathbf{Z} \in R_i) = \alpha$ and (ii) for every $j = 1, \ldots, K$, $j \neq i$, $\lim_{l \to \infty} P_{\theta_l}(\mathbf{Z} \in R_j) = 1$ where $\mathbf{Z}$ is the data.

Let $\Theta_{\mathbf{U}^T\mathbf{V}} = \{\Theta_{\mathbf{U}^T\mathbf{V}}^A, \Theta_{\mathbf{U}^T\mathbf{V}}^B\}$, $\Theta_{\mathbf{U}^T\mathbf{V}}^A = \bigcup_{k=1,\ldots,2^{p-1}} \Theta_{\mathbf{U}^T\mathbf{V},k}^A$, and $\Theta_{\mathbf{U}^T\mathbf{V}}^B = \bigcup_{k=1,\ldots,2^{p-1}} \Theta_{\mathbf{U}^T\mathbf{V},k}^B$, where $\Theta_{\mathbf{U}^T\mathbf{V},k}^A = \{\boldsymbol{\theta} : w_1\alpha_{S1} = 0, \sum_{j=2}^{p} w_j\theta_j^* = 0\}$ and $\Theta_{\mathbf{U}^T\mathbf{V},k}^B = \{\boldsymbol{\theta} : w_1\beta_{M1} = 0, \sum_{j=2}^{p} w_j\theta_j^* = 0\}$. For JT of $\mathbf{U}_n^T\mathbf{V}_n$, suppose that $\theta_l = (\boldsymbol{\alpha}_{Sl}^T, \boldsymbol{\beta}_{Ml}^T)^T = (\mathbf{0}^T, \boldsymbol{\beta}_{Ml}^T)^T$ where all elements in $\boldsymbol{\beta}_{Ml}$ except β_{M1} go to 0 and $\lim_{l \to \infty} \beta_{M1l} \to \infty$, and thus $\theta_l \in \Theta_{\mathbf{U}^T\mathbf{V}}^A$. Denote $\sigma_{\alpha 1}^2$ and $\sigma_{\beta 1}^2$ as variances of $\sqrt{n}\hat{\alpha}_{S1}$ and $\sqrt{n}\hat{\beta}_{M1}$, respectively, and $\Phi(z_{1-\alpha}) = 1 - \alpha$. One can show that

$$\lim_{l \to \infty} P_{\theta_l}(\mathbf{Z} \in R_{\mathbf{U}^T\mathbf{V}}^A) = \lim_{l \to \infty} P_{\theta_l}\left(\left|\frac{w_1\hat{\alpha}_{S1l} + \sum_{j=2}^{p} w_j\hat{\theta}_j^*}{\sqrt{V_{\max}}}\right| > z_{1-\alpha} \,\Big|\, \alpha_{S1l} = 0\right)$$

$$= \lim_{l \to \infty} P_{\theta_l}\left(\left|\frac{\sqrt{n}\hat{\beta}_{M1l}\hat{\alpha}_{S1l}}{\sqrt{\hat{\beta}_{M1l}^2 \sigma_{\alpha 1}}} + o_p(1)\right| > z_{1-\alpha} \,\Big|\, \alpha_{S1l} = 0\right)$$

$$\to \alpha$$

and

$$\lim_{l \to \infty} P_{\theta_l}(\mathbf{Z} \in R_{\mathbf{U}^T\mathbf{V}}^B) = \lim_{l \to \infty} P_{\theta_l}\left(\left|\frac{w_1\hat{\beta}_{M1l} + \sum_{j=2}^{p} w_j\hat{\theta}_j^*}{\sqrt{V_{\max}}}\right| > z_{1-\alpha} \,\Big|\, \beta_{M1l} \neq 0\right)$$

$$= \lim_{l \to \infty} P_{\theta_l}\left(\left|\frac{\sqrt{n}w_1\hat{\beta}_{M1l} + O_p(1)}{\sqrt{nV_{\max}}}\right| > z_{1-\alpha} \,\Big|\, \beta_{M1l} \neq 0\right)$$

$$\to 1,$$

because $nV_{\max}$ is bounded, $\sqrt{n}w_1 = O_p(1)$ and $\hat{\beta}_{M1l} \to \infty$.

For JT of $U_nV_nW_n$, suppose that $\theta_l = (\gamma_{Sl}, \alpha_{Ml}, \beta_{Gl})^T = (0, \alpha_{Ml}, \beta_{Gl})^T \in \Theta_{UVW1}$, $\lim_{l \to \infty} \alpha_{Ml} = \infty$, and $\lim_{l \to \infty} \beta_{Gl} = \infty$. It can be shown that

$$\lim_{l \to \infty} P_{\theta_l}(\mathbf{Z} \in R_{UVW1}) = \lim_{l \to \infty} P_{\theta_l}\left(\left|\frac{\sqrt{n}\hat{\gamma}_{Sl}}{\sqrt{n\sigma_{\gamma n}^2}}\right| > z_{1-\alpha} \,\Big|\, \gamma_{Sl} = 0\right) \to \alpha,$$

$$\lim_{l \to \infty} P_{\theta_l}(\mathbf{Z} \in R_{UVW2}) = \lim_{l \to \infty} P_{\theta_l}\left(\left|\frac{\sqrt{n}\hat{\alpha}_{Ml}}{\sqrt{n\sigma_{\alpha n}^2}}\right| > z_{1-\alpha} \mid \alpha_{Ml} \neq 0\right) \to 1,$$

and similarly, $\lim_{l \to \infty} P_{\theta_l}(\mathbf{Z} \in R_{UVW3}) = \lim_{l \to \infty} P_{\theta_l}(|\frac{\sqrt{n}\hat{\beta}_{Gl}}{\sqrt{n\sigma_{\beta n}^2}}| > z_{1-\alpha} \mid \beta_{Gl} \neq 0) \to 1$. Therefore, both JTs are IUT and satisfy the two conditions (i) and (ii). By Theorem 2 in Berger and Hsu (1996), they are size α tests.   □

SUPPLEMENTARY MATERIAL

Supplement to "Joint significance tests for mediation effects of socioeconomic adversity on adiposity via epigenetics" (DOI: 10.1214/17-AOAS1120SUPP; .pdf). Supplementary material includes discussion of causal assumptions, additional simulation studies, and PSE analyses of 26 methylation loci of *FASN*.

REFERENCES

AGHA, G., HOUSEMAN, E. A., KELSEY, K. T., EATON, C. B., BUKA, S. L. and LOUCKS, E. B. (2015). Adiposity is associated with DNA methylation profile in adipose tissue. *Int. J. Epidemiol.* **44** 1277–1287.

ALBERT, J. M. and NELSON, S. (2011). Generalized causal mediation analysis. *Biometrics* **67** 1028–1038. MR2829237

AROIAN, L. A. (1947). The probability function of the product of two normally distributed variables. *Ann. Math. Stat.* **18** 265–271. MR0021284

AVIN, C., SHPITSER, I. and PEARL, J. (2005). Identifiability of path-specific effects. In *Proceedings of the International Joint Conferences on Artificial Intelligence* 357–363.

BARON, R. M. and KENNY, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical consideration. *J. Pers. Soc. Psychol.* **51** 1173–1182.

BERGER, R. L. and HSU, J. C. (1996). Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statist. Sci.* **11** 283–319. MR1445984

BOLLEN, K. A. and STINE, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociol. Method.* **20** 115–140.

BORGHOL, N., SUDERMAN, M., MCARDLE, W., RACINE, A., HALLETT, M., PEMBREY, M., HERTZMAN, C., POWER, C. and SZYF, M. (2012). Associations with early-life socioeconomic position in adult DNA methylation. *Int. J. Epidemiol.* **41** 62–74.

DARMON, N. and DREWNOWSKI, A. (2008). Does social class predict diet quality? *Am. J. Clin. Nutr.* **87** 1107–1117.

DAVIS, S., DU, P., BILKE, S., TRICHE, T. J. and BOOTWALLA, M. (2015). methylumi: Handle Illumina methylation data. R package version 2.14.0.

GISKES, K., AVENDANO, M., BRUG, J. and KUNST, A. E. (2010). A systematic review of studies on socioeconomic inequalities in dietary intakes associated with weight gain and overweight/obesity conducted among European adults. *Obes. Rev.* **11** 413–429.

HARDY, J. B. (1971). The Johns Hopkins collaborative perinatal project. Descriptive background. *Johns Hopkins Med. J.* **128** 238–243.

HUANG, Y.-T. (2018). Supplement to "Joint significance tests for mediation effects of socioeconomic adversity on adiposity via epigenetics." DOI:10.1214/17-AOAS1120SUPP.

HUANG, Y.-T. and CAI, T. (2016). Mediation analysis for survival data using semiparametric probit models. *Biometrics* **72** 563–574. MR3515783

HUANG, Y.-T. and PAN, W.-C. (2016). Hypothesis test of mediation effect in causal mediation model with high-dimensional continuous mediators. *Biometrics* **72** 401–413. MR3515767

HUANG, Y. T., CHU, S., LOUCKS, E. B., LIN, C. L., EATON, C. B., BUKA, S. L. and KELSEY, K. T. (2016). Epigenome-wide profiling of DNA methylation in paired samples of adipose tissue and blood. *Epigenetics* **11** 227–236.

IMAI, K., KEELE, L. and YAMAMOTO, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statist. Sci.* **25** 51–71. MR2741814

KADOTA, Y., KAWAKAMI, T., TAKASAKI, S., SATO, M. and SUZUKI, S. (2016). Gene expression related to lipid and glucose metabolism in white adipose tissue. *Obes. Res. Clin. Pract.* **10** 85–93.

KOVACS, P., HARPER, I., HANSON, R. I., INFANTE, A. M., BOGARDUS, C., TATARANNI, P. A. and BAIER, L. J. (2004). A novel missense substitution (Val1483Ile) in the fatty acid synthase gene (FAS) is associated with percentage of body fat and substrate oxidation rates in nondiabetic Pima Indians. *Diabetes* **53** 1915–1919.

LANGE, T. and HANSEN, J. V. (2011). Direct and indirect effects in a survival context. *Epidemiology* **22** 575–581.

LOFTUS, T. M., JAWORSKY, D. E., FREHYWOT, G. L., TOWNSEND, C. A., RONNETT, G. V., LANE, M. D. and KUHAJDA, F. P. (2000). Reduced food intake and body weight in mice treated with fatty acid synthase inhibitors. *Science* **288** 2379–2381.

LOUCKS, E. B., HUANG, Y. T., AGHA, G., CHU, S., EATON, C. B., GILMAN, S. E., BUKA, S. L. and KELSEY, K. T. (2016). Epigenetic mediators between childhood socioeconomic disadvantage and mid-life body mass index: The New England Family Study. *Psychosom. Med.* **78** 1053–1065.

MACKINNON, D. (2008). *Introduction to Statistical Mediation Analysis*. Taylor & Francis, London.

MACKINNON, D. P., LOCKWOOD, C. M., HOFFMAN, J. M., WEST, S. G. and SHEETS, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychol. Methods* **7** 83–104.

PEARL, J. (2001). Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence* 411–420. Morgan Kaufmann, San Francisco.

PREACHER, K. J. and HAYES, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* **40** 879–891.

ROBINS, J. M. (2003). *Semantics of Causal DAG Models and the Identification of Direct and Indirect Effects*. Oxford Univ. Press, New York.

ROBINS, J. M. and GREENLAND, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology* **3** 143–155.

RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58. MR0472152

SENESE, L. C., ALMEIDA, N. D., FATH, A. K., SMITH, B. T. and LOUCKS, E. B. (2009). Associations between childhood socioeconomic position and adulthood obesity. *Epidemiol. Rev.* **31** 21–51.

SOBEL, M. E. (1982). *Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models*. American Sociological Association, Washington, DC.

TAYLOR, A. B., MACKINNON, D. P. and TEIN, J.-Y. (2008). Tests of the three-path mediated effect. *Organ. Res. Methods* **11** 241–269.

TCHETGEN TCHETGEN, E. J. (2011). On causal mediation analysis with a survival outcome. *Int. J. Biostat.* **7** Article 33. MR2843528

TESCHENDORFF, A. E., MARABITA, F., LECHNER, M., BARTLETT, T., TEGNER, J., GOMEZ-CABRERO, D. and BECK, S. (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29** 189–196.

VANDERWEELE, T. J. (2011). Causal mediation analysis with survival data. *Epidemiology* **22** 582–585.

VANDERWEELE, T. J. and VANSTEELANDT, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Stat. Interface* **2** 457–468. MR2576399

VANDERWEELE, T. J. and VANSTEELANDT, S. (2010). Odds ratios for mediation analysis for a dichotomous outcome. *Am. J. Epidemiol.* **172** 1339–1348.

VANDERWEELE, T. J. and VANSTEELANDT, S. (2013). Mediation analysis with multiple mediators. *Epidemiol. Methods* **2** 95–115.

WU, M., KRAFT, P., EPSTEIN, M., TAYLOR, D., CHANOCK, S., HUNTER, D. J. and LIN, X. (2010). Powerful SNP set analysis for case-control genomewide association studies. *Am. J. Hum. Genet.* **86** 929–942.

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
128 ACADEMIA ROAD, SECTION 2
TAIPEI 11529
TAIWAN
E-MAIL: ythuang@stat.sinica.edu.tw