Biometrika (2014), **xx**, x, *pp*. 1–15

Printed in Great Britain

Supplementary material for Using covariate-specific disease prevalence information to increase the power of case-control studies

BY JING QIN

National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892

jingqin@niaid.nih.gov

HAN ZHANG

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892

han.zhang2@nih.gov

PENGFEI LI

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

pengfei.li@uwaterloo.ca

DEMETRIUS ALBANES AND KAI YU

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892

albanesd@mail.nih.gov yuka@mail.nih.gov

This is a supplementary document to the corresponding paper submitted to *Biometrika*. We summarize the main results in Section 1 and present some preliminary preparation in Section 2. The proofs of Theorem 1, Corollary 1, and Theorem 2 are given in Sections 3, 4, and 5, respectively. More simulation results are provided in Section 6.

1. MAIN RESULTS

THEOREM 1. *Suppose $\rho = n_1/n_0$ remains constant as $n \to \infty$ and $\rho \in (0,1)$. Under regularity conditions, as n goes to infinity,*

$$n^{1/2}(\hat{\theta} - \theta_0) \to N(0, \Sigma)$$

in distribution, where $\Sigma = J^{-1} - \rho^{-1}(1+\rho)^2(1,1,0,0,0)^{\mathrm{T}}(1,1,0,0,0)$ and $J = UV^{-1}U^{\mathrm{T}}$. The matrices U and V are defined in (A4) *of the main article.*

COROLLARY 1. *Let $\theta_{02} = (\beta_0, \gamma_0, \xi_0)^{\mathrm{T}}$ be the true value of θ_2 and let $\hat{\theta}_{2L}$ be the maximum likelihood estimator of θ_2 based on logistic regression in the absence of auxiliary information. Under the conditions of Theorem* 1*, we have:*

(a) *if $I = 1$, the asymptotic variance of $n^{1/2}(\hat{\theta}_2 - \theta_{02})$ is the same as that of $n^{1/2}(\hat{\theta}_{2L} - \theta_{02})$;*

(b) *if* $I > 1$, *the difference of the asymptotic covariance matrices of* $n^{1/2}(\hat{\theta}_2 - \theta_{02})$ *and* $n^{1/2}(\hat{\theta}_{2L} - \theta_{02})$ *is non-positive-definite; and*

(c) *when* $I > 1$, *the asymptotic covariance matrix of* $n^{1/2}(\hat{\theta}_2 - \theta_{02})$ *can not decrease if an estimating equation in* (4) *of the main article is dropped.*

THEOREM 2. *Under the conditions of Theorem* 1, *as* n *goes to infinity, the empirical likelihood ratio statistic* $R(0) \to \chi^2_1$ *in distribution if* $\xi = 0$.

2. PRELIMINARY PREPARATION

We need some notation. Let $Z = (X, Y, XY)^{\mathrm{T}}$ and $Z_i = (X_i, Y_i, X_iY_i)^{\mathrm{T}}$. Then

$$\delta(X, Y; \theta) = \exp(\alpha^* + \theta_2^{\mathrm{T}} Z).$$

For the auxiliary information, recall that $g(X, Y; \theta) = \{g_1(X, Y; \theta), \ldots, g_I(X, Y; \theta)\}^{\mathrm{T}}$. Further,

$$E_0\{\delta(X, Y; \theta) - 1\} = 0, \;\; E_0\{g(X, Y; \theta)\} = 0.$$

Here $E_0(\cdot)$ means taking the expectation under $f(x, y \mid D = 0)$, the joint distribution of (X, Y) under the control group. Similarly, $E_1(\cdot)$ means taking the expectation under $f(x, y \mid D = 1)$, the joint distribution of (X, Y) under the case group.

With the notation introduced above, the profile likelihood $\ell(\theta)$ can be written

$$\ell(\theta) = \sum_{i=1}^{n} D_i(\alpha^* + \theta_2^{\mathrm{T}} Z_i) - \sum_{i=1}^{n} \log[1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)]$$

with the Lagrange multipliers λ and t being determined by

$$\sum_{i=1}^{n} \frac{\delta(X_i, Y_i; \theta) - 1}{1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)} = 0,$$

$$\sum_{i=1}^{n} \frac{g(X_i, Y_i; \theta)}{1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)} = 0.$$

The true values of λ and t are $\lambda_0 = n_1/n$ and 0, respectively.

We note that the profile likelihood $\ell(\theta)$ can be written as $\ell(\theta) = \inf_{t,\lambda} l(\theta, t, \lambda)$ with

$$l(\theta, t, \lambda) = \sum_{i=1}^{n} D_i(\alpha^* + \theta_2^{\mathrm{T}} Z_i) - \sum_{i=1}^{n} \log[1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)].$$

Equivalently, $\ell(\theta) = l(\theta, t, \lambda)$ with t and λ being the solution to $\partial l(\theta, t, \lambda)/\partial t = 0$ and $\partial l(\theta, t, \lambda)/\partial \lambda = 0$.

To investigate the asymptotic properties of $\hat{\theta}$, we need its approximation, which can be obtained via the second-order Taylor expansion on $l(\theta, t, \lambda)$.

2·1. *First derivatives of* $l(\theta, t, \lambda)$

We first calculate the first derivatives of $l(\theta, t, \lambda)$, which are as follows:

$$\frac{\partial l(\theta,t,\lambda)}{\partial \eta} = -\sum_{i=1}^{n} \frac{t^{\mathrm{T}} \partial g(X_i, Y_i; \theta)/\partial \eta}{1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)}, \tag{1}$$

$$\frac{\partial l(\theta,t,\lambda)}{\partial \alpha^*} = \sum_{i=1}^{n} D_i - \sum_{i=1}^{n} \frac{\lambda\delta(X_i, Y_i; \theta) + t^{\mathrm{T}} \partial g(X_i, Y_i; \theta)/\partial \alpha^*}{1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)}, \tag{2}$$

$$\frac{\partial l(\theta,t,\lambda)}{\partial \theta_2} = \sum_{i=1}^{n} D_i Z_i - \sum_{i=1}^{n} \frac{\lambda\delta(X_i, Y_i; \theta) Z_i + \{\partial g^{\mathrm{T}}(X_i, Y_i; \theta)/\partial \theta_2\} t}{1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)}, \tag{3}$$

$$\frac{\partial l(\theta,t,\lambda)}{\partial \lambda} = -\sum_{i=1}^{n} \frac{\delta(X_i, Y_i; \theta) - 1}{1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)}, \tag{4}$$

$$\frac{\partial l(\theta,t,\lambda)}{\partial t} = -\sum_{i=1}^{n} \frac{g(X_i, Y_i; \theta)}{1 + \lambda\{\delta(X_i, Y_i; \theta) - 1\} + t^{\mathrm{T}} g(X_i, Y_i; \theta)}. \tag{5}$$

Setting (1)–(2) and (4)–(5) equal to 0 at $\theta = \hat{\theta}$, $t = \hat{t}$, and $\lambda = \hat{\lambda}$ and noting that $\partial g(X, Y; \theta)/\partial \eta = -\partial g(X, Y; \theta)/\partial \alpha^*$, we get

$$\hat{\lambda} = n_1/n = \lambda_0. \tag{6}$$

Recall that $\rho = n_1/n_0$, $\omega = (\theta^{\mathrm{T}}, t^{\mathrm{T}})^{\mathrm{T}}$, and $\omega_0 = (\theta_0^{\mathrm{T}}, 0)^{\mathrm{T}}$ is the true value of ω. Let $\delta_i = \delta(X_i, Y_i; \theta_0)$ and $\Delta_i = 1 + \rho\delta_i$. Then

$$\frac{\partial l(\theta_0, 0, \lambda_0)}{\partial \omega} = \begin{pmatrix} \frac{\partial l(\theta_0,0,\lambda_0)}{\partial \eta} \\ \frac{\partial l(\theta_0,0,\lambda_0)}{\partial \alpha^*} \\ \frac{\partial l(\theta_0,0,\lambda_0)}{\partial \theta_2} \\ \frac{\partial l(\theta_0,0,\lambda_0)}{\partial t} \end{pmatrix} = \begin{pmatrix} 0 \\ S_{n2} \\ S_{n3} \\ S_{n4} \end{pmatrix} \tag{7}$$

with

$$S_{n2} = n_1 - \rho \sum_{i=1}^{n} \frac{\delta_i}{\Delta_i}, \quad S_{n3} = \sum_{i=1}^{n} D_i Z_i - \rho \sum_{i=1}^{n} \frac{\delta_i Z_i}{\Delta_i}, \quad S_{n4} = -(1+\rho) \sum_{i=1}^{n} \frac{g(X_i, Y_i; \theta_0)}{\Delta_i}.$$

Further, let

$$S_n = (S_{n2}, S_{n3}^{\mathrm{T}}, S_{n4}^{\mathrm{T}})^{\mathrm{T}}. \tag{8}$$

2·2. *Second derivatives of* $l(\theta, t, \lambda)$

We next calculate the second derivatives of $l(\theta, t, \lambda)$ with respect to θ and t. Since $\hat{\lambda} = \lambda_0$, in the following derivation we set λ to λ_0. After some calculation, it can be verified that the second derivatives of $l(\theta, t, \lambda_0)$ at $\omega = \omega_0$ are

$$\frac{\partial^2 l(\theta_0, 0, \lambda_0)}{\partial \omega \partial \omega^{\mathrm{T}}} = \begin{pmatrix} \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \eta^2} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \eta \partial \alpha^*} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \eta \partial \theta_2^{\mathrm{T}}} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \eta \partial t^{\mathrm{T}}} \\ \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \eta \partial \alpha^*} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial (\alpha^*)^2} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \alpha^* \partial \theta_2^{\mathrm{T}}} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \alpha^* \partial t^{\mathrm{T}}} \\ \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \eta \partial \theta_2} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \alpha^* \partial \theta_2} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \theta_2 \partial \theta_2^{\mathrm{T}}} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \theta_2 \partial t^{\mathrm{T}}} \\ \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \eta \partial t} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial \alpha^* \partial t} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial t \partial \theta_2^{\mathrm{T}}} & \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial t \partial t^{\mathrm{T}}} \end{pmatrix} \tag{9}$$

with

$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\eta^2}=0,$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\eta\partial\alpha^*}=0,$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\eta\partial\theta_2^{\mathrm{T}}}=0,$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\eta\partial t^{\mathrm{T}}}=-\sum_{i=1}^{n}\frac{(1+\rho)\partial g^{\mathrm{T}}(X_i,Y_i;\theta_0)/\partial\eta}{\Delta_i},$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial(\alpha^*)^2}=-\sum_{i=1}^{n}\frac{\rho\delta_i}{\Delta_i^2},$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\alpha^*\partial\theta_2^{\mathrm{T}}}=-\sum_{i=1}^{n}\frac{\rho\delta_i Z_i^{\mathrm{T}}}{\Delta_i^2},$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\alpha^*\partial t^{\mathrm{T}}}=(1+\rho)\sum_{i=1}^{n}\frac{\partial g^{\mathrm{T}}(X_i,Y_i;\theta_0)/\partial\eta}{\Delta_i}+\rho(1+\rho)\sum_{i=1}^{n}\frac{\delta_i g^{\tau}(X_i,Y_i;\theta_0)}{\Delta_i^2},$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\theta_2\partial\theta_2^{\mathrm{T}}}=-\sum_{i=1}^{n}\frac{\rho\delta_i Z_i Z_i^{\mathrm{T}}}{\Delta_i^2},$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\theta_2\partial t^{\mathrm{T}}}=-(1+\rho)\sum_{i=1}^{n}\frac{\partial g^{\mathrm{T}}(X_i,Y_i;\theta_0)/\partial\theta_2}{\Delta_i}+\rho(1+\rho)\sum_{i=1}^{n}\frac{\delta_i Z_i g^{\mathrm{T}}(X_i,Y_i;\theta_0)}{\Delta_i},$$
$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial t\partial t^{\mathrm{T}}}=(1+\rho)^2\sum_{i=1}^{n}\frac{g(X_i,Y_i;\theta_0)g^{\mathrm{T}}(X_i,Y_i;\theta_0)}{\Delta_i^2}.$$

2·3. *Some useful technical lemmas*

When deriving the asymptotic distribution of $\hat{\theta}$, we need to use $E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}\right\}$ and the expectation and variance of S_n defined in (8). We need the following lemma to simplify our calculation.

LEMMA 1. *Let $h(Z)$ be an arbitrary function of Z. Further, let $\delta_0=\delta(X,Y;\theta_0)$ and $\Delta_0=1+\rho\delta_0$. Then we have*

$$E_1\{H(Z)\}=E_0\{\delta_0 H(Z)\}$$

and

$$E\left\{\sum_{i=1}^{n}h(Z_i)\right\}=\frac{n}{1+\rho}E_0\left\{\Delta_0 h(Z)\right\}.$$

Proof. Let $z=(x,y,xy)^{\mathrm{T}}$. Then

$$E_1\{H(Z)\}=\int h(z)dF_1(x,y).$$

With $dF_1(x,y) = \delta(x,y;\theta_0)dF_0(x,y)$, we have that

$$E_1\{H(Z)\} = \int \delta(x,y;\theta_0)h(z)dF_0(x,y) = E\{\delta_0 h(Z)\}.$$

This completes the proof of the first part.

For the second part, we have

$$E\left\{\sum_{i=1}^{n} h(Z_i)\right\} = n_1 E_1\{h(Z)\} + n_0 E_0\{h(Z)\} = n_1 E\{\delta_0 h(Z)\} + n_0 E_0\{h(Z)\}.$$

Further, $\rho = n_1/n_0$ implies that $n_0 = n/(1+\rho)$ and $n_1 = n\rho/(1+\rho)$. Therefore,

$$E\left\{\sum_{i=1}^{n} h(Z_i)\right\} = \frac{n}{1+\rho} E_0\{(\rho\delta_0 + 1)h(Z)\} = \frac{n}{1+\rho} E_0\left\{\Delta_0 h(Z)\right\}.$$

This completes the proof of the second part. □

With the help of Lemma 1, we find $E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}\right\}$ in the next lemma.

LEMMA 2. *With* $E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}\right\}$ *defined in* (9)*, we have*

$$\frac{1}{n} E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}\right\} = \begin{pmatrix} 0 & 0 & 0 & -A_{14} \\ 0 & -A_{22} & -A_{23} & -A_{24} \\ 0 & -A_{32} & -A_{33} & -A_{34} \\ -A_{41} & -A_{42} & -A_{43} & A_{44} \end{pmatrix}$$

with

$$A_{14} = A_{41}^{\mathrm{T}} = E_0\left\{\frac{\partial g^{\mathrm{T}}(X,Y;\theta_0)}{\partial\eta}\right\}, \tag{10}$$

$$A_{22} = \frac{\rho}{1+\rho} E_0\left(\frac{\delta_0}{\Delta_0}\right), \tag{11}$$

$$A_{23} = A_{32}^{\mathrm{T}} = \frac{\rho}{1+\rho} E_0\left(\frac{\delta_0 Z^{\mathrm{T}}}{\Delta_0}\right), \tag{12}$$

$$A_{24} = A_{42}^{\mathrm{T}} = -E_0\left\{\frac{\partial g^{\mathrm{T}}(X,Y;\theta_0)}{\partial\eta}\right\} + E_0\left\{\frac{g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0}\right\}, \tag{13}$$

$$A_{33} = \frac{\rho}{1+\rho} E_0\left(\frac{\delta_0 Z Z^{\mathrm{T}}}{\Delta_0}\right), \tag{14}$$

$$A_{34} = A_{43}^{\mathrm{T}} = E_0\left\{\frac{\partial g^{\mathrm{T}}(X,Y;\theta_0)}{\partial\theta_2}\right\} - \rho\left\{\frac{\delta_0 Z g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0}\right\}, \tag{15}$$

$$A_{44} = (1+\rho) E_0\left\{\frac{g(X,Y;\theta_0) g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0}\right\}. \tag{16}$$

Proof. We verify only that

$$\frac{1}{n} E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\alpha^*\partial t^{\mathrm{T}}}\right\} = -A_{24} = E_0\left\{\frac{\partial g^{\mathrm{T}}(X,Y;\theta_0)}{\partial\eta}\right\} - E_0\left\{\frac{g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0}\right\}. \tag{17}$$

For the other parts, the procedure and idea are similar and are omitted.

Recall that

$$\frac{\partial^2 l(\theta_0, 0, \lambda_0)}{\partial \alpha^* \partial t^{\mathrm{T}}} = (1+\rho)\sum_{i=1}^{n} \frac{\partial g^{\mathrm{T}}(X_i, Y_i; \theta_0)/\partial \eta}{\Delta_i} + \rho(1+\rho)\sum_{i=1}^{n} \frac{\delta_i g^{\tau}(X_i, Y_i; \theta_0)}{\Delta_i^2}.$$

Then

$$E\left\{\frac{\partial^2 l(\theta_0, 0, \lambda_0)}{\partial \alpha^* \partial t^{\mathrm{T}}}\right\} = (1+\rho)E\left\{\sum_{i=1}^{n} \frac{\partial g^{\mathrm{T}}(X_i, Y_i; \theta_0)/\partial \eta}{\Delta_i}\right\} + \rho(1+\rho)E\left\{\sum_{i=1}^{n} \frac{\delta_i g^{\tau}(X_i, Y_i; \theta_0)}{\Delta_i^2}\right\}. \tag{18}$$

Applying Lemma 1 to the two terms on the right-hand side of (18), we have

$$(1+\rho)E\left\{\sum_{i=1}^{n} \frac{\partial g^{\mathrm{T}}(X_i, Y_i; \theta_0)/\partial \eta}{\Delta_i}\right\} = nE_0\left\{\frac{\partial g^{\mathrm{T}}(X, Y; \theta_0)}{\partial \eta}\right\} \tag{19}$$

and

$$\rho(1+\rho)E\left\{\sum_{i=1}^{n} \frac{\delta_i g^{\tau}(X_i, Y_i; \theta_0)}{\Delta_i^2}\right\} = n\rho E_0\left\{\frac{\delta_0 g^{\tau}(X, Y; \theta_0)}{\Delta_0}\right\}. \tag{20}$$

Recall that $\Delta_0 = 1 + \rho\delta_0$. Then

$$\delta_0 = \frac{\Delta_0 - 1}{\rho}$$

and

$$E_0\left\{\frac{\delta_0 g(X, Y; \theta_0)}{\Delta_0}\right\} = \frac{1}{\rho}\left[E_0\left\{g(X, Y; \theta_0)\right\} - E_0\left\{\frac{g(X, Y; \theta_0)}{\Delta_0}\right\}\right] = -\frac{1}{\rho}E_0\left\{\frac{g(X, Y; \theta_0)}{\Delta_0}\right\}. \tag{21}$$

The last step in (21) follows from the fact that $E_0\left\{g(X, Y; \theta_0)\right\} = 0$.

Substituting (21) into (20) gives

$$\rho(1+\rho)E\left\{\sum_{i=1}^{n} \frac{\delta_i g^{\tau}(X_i, Y_i; \theta_0)}{\Delta_i^2}\right\} = -nE_0\left\{\frac{g^{\mathrm{T}}(X, Y; \theta_0)}{\Delta_0}\right\}. \tag{22}$$

Combining (18), (19), and (22), we verify that (17) is correct. This completes the proof of Lemma 2. □

The final lemma presents the expectation and variance of S_n.

LEMMA 3. *With S_n defined in* (8)*, we have*

$$E(S_n) = 0$$

and

$$\frac{1}{n}\mathrm{var}(S_n) = \begin{pmatrix} A_{22} & A_{23} & 0 \\ A_{32} & A_{33} & 0 \\ 0 & 0 & A_{44} \end{pmatrix} - \frac{(1+\rho)^2}{\rho}\begin{pmatrix} A_{22} \\ A_{32} \\ A_{42}+A_{41} \end{pmatrix}\begin{pmatrix} A_{22} \\ A_{32} \\ A_{42}+A_{41} \end{pmatrix}^{\mathrm{T}}.$$

Proof. For $E(S_n)$, we show only that $E(S_{n3}) = 0$. The other parts, $E(S_{n2}) = 0$ and $E(S_{n4}) = 0$, can be verified using Lemma 1 directly. Recall that

$$S_{n3} = \sum_{i=1}^{n} D_i Z_i - \rho \sum_{i=1}^{n} \frac{\delta_i Z_i}{\Delta_i}.$$

Applying Lemma 1 and noting that $n_1 = n\rho/(1+\rho)$, we have that

$$E(S_{n3}) = n_1 E_1(Z) - \frac{n\rho}{1+\rho} E(\delta_0 Z) = \frac{n\rho}{1+\rho} E_0(\delta_0 Z) - \frac{n\rho}{1+\rho} E(\delta_0 Z) = 0.$$

For $n^{-1}\mathrm{var}(S_n)$, we verify only that

$$\mathrm{var}(S_{n4}) = nA_{44} - \frac{n(1+\rho)^2}{\rho}(A_{42} + A_{41})(A_{42} + A_{41})^{\mathrm{T}}. \tag{23}$$

The other parts, again, can be similarly checked.

Recall that

$$S_{n4} = -(1+\rho) \sum_{i=1}^{n} \frac{g(X_i, Y_i; \theta_0)}{\Delta_i}.$$

Then

$$\mathrm{var}(S_{n4}) = n_1(1+\rho)^2 \mathrm{var}_1 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\} + n_0(1+\rho)^2 \mathrm{var}_0 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\}, \tag{24}$$

where $\mathrm{var}_1(\cdot)$ and $\mathrm{var}_0(\cdot)$ mean that the variances are calculated under $f(x, y \mid D = 1)$ and $f(x, y \mid D = 0)$, respectively.

Applying Lemma 1, we have

$$\begin{aligned}
&\mathrm{var}_1 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\} \\
&= E_1 \left\{ \frac{g(X,Y;\theta_0) g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0^2} \right\} - E_1 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\} E_1 \left\{ \frac{g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0} \right\} \\
&= E_0 \left\{ \frac{\delta_0 g(X,Y;\theta_0) g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0^2} \right\} - E_0 \left\{ \frac{\delta_0 g(X,Y;\theta_0)}{\Delta_0} \right\} E_0 \left\{ \frac{\delta_0 g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0} \right\} \\
&= E_0 \left\{ \frac{\delta_0 g(X,Y;\theta_0) g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0^2} \right\} - \frac{1}{\rho^2} E_0 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\} E_0 \left\{ \frac{g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0} \right\},
\end{aligned} \tag{25}$$

where we have used the result in (21).

Similarly, we get

$$\mathrm{var}_0 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\} = E_0 \left\{ \frac{g(X,Y;\theta_0) g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0^2} \right\} - E_0 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\} E_0 \left\{ \frac{g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0} \right\}. \tag{26}$$

Combining (24)–(26), we obtain

$$\begin{aligned}
\mathrm{var}(S_{n4}) = {} & E_0 \left\{ \frac{(n_1\delta_0 + n_0)(1+\rho)^2 g(X,Y;\theta_0) g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0^2} \right\} \\
& -(n_1/\rho^2 + n_0)(1+\rho)^2 E_0 \left\{ \frac{g(X,Y;\theta_0)}{\Delta_0} \right\} E_0 \left\{ \frac{g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0} \right\}.
\end{aligned}$$

Since $n_1 = n\rho/(1+\rho)$, $n_0 = n/(1+\rho)$, and $\Delta_0 = 1 + \rho\delta_0$, we can simplify $\mathrm{var}(S_{n4})$ to

$$\mathrm{var}(S_{n4}) = n(1+\rho)E_0\left\{\frac{g(X,Y;\theta_0)g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0^2}\right\} - \frac{n(1+\rho)^2}{\rho}E_0\left\{\frac{g(X,Y;\theta_0)}{\Delta_0}\right\}E_0\left\{\frac{g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0}\right\}.$$

From (10)–(16), we notice that

$$n(1+\rho)E_0\left\{\frac{g(X,Y;\theta_0)g^{\mathrm{T}}(X,Y;\theta_0)}{\Delta_0^2}\right\} = nA_{44}$$

and

$$E_0\left\{\frac{g(X,Y;\theta_0)}{\Delta_0}\right\} = A_{41} + A_{42}.$$

Hence,

$$\mathrm{var}(S_{n4}) = nA_{44} - \frac{n(1+\rho)^2}{\rho}(A_{41}+A_{42})(A_{41}+A_{42})^{\mathrm{T}}$$

as claimed in (23). This completes the proof of Lemma 3. □

3. PROOF OF THEOREM 1

Using a similar argument to that used in the proofs of Lemma 1 and Theorem 1 of Qin & Lawless (1994), we have that $\hat{\theta} = \theta_0 + O_p(n^{-1/2})$ and $\hat{t} = O_p(n^{-1/2})$. Next we investigate the asymptotic approximation of $\hat{\theta}$.

The maximum likelihood estimator $\hat{\theta}$ of θ and the associated Lagrange multiplier $\hat{t}$ must satisfy

$$\frac{\partial l(\hat{\theta},\hat{t},\hat{\lambda})}{\partial\omega} = 0.$$

Recall that $\hat{\lambda} = \lambda_0$. Then

$$\frac{\partial l(\hat{\theta},\hat{t},\lambda_0)}{\partial\omega} = 0.$$

Applying a first-order expansion to $\frac{\partial l(\hat{\theta},\hat{t},\lambda_0)}{\partial\omega}$ gives

$$0 = \frac{\partial l(\theta_0,0,\lambda_0)}{\partial\omega} + \frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}(\hat{\omega}-\omega_0) + o_p(n^{1/2}). \tag{27}$$

The law of large numbers implies that

$$\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}} = E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}\right\} + o_p(n). \tag{28}$$

Combining (27) and (28), we get

$$E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}\right\}(\hat{\omega}-\omega_0) = -\frac{\partial l(\theta_0,0,\lambda_0)}{\partial\omega} + o_p(n^{1/2}). \tag{29}$$

Using the forms of $E\left\{\frac{\partial^2 l(\theta_0,0,\lambda_0)}{\partial\omega\partial\omega^{\mathrm{T}}}\right\}$ and $\frac{\partial l(\theta_0,0,\lambda_0)}{\partial\omega}$ provided in Lemma 2 and Equation (7), respectively, we have

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -A_{22} & -A_{23} \\ 0 & -A_{32} & -A_{33} \end{pmatrix} (\hat{\theta}-\theta_0) - \begin{pmatrix} A_{14} \\ A_{24} \\ A_{34} \end{pmatrix} \hat{t} = -n^{-1} \begin{pmatrix} 0 \\ S_{n2} \\ S_{n3} \end{pmatrix} + o_p(n^{-1/2}), \tag{30}$$

$$-\ (A_{41}, A_{42}, A_{43})\,(\hat{\theta}-\theta_0) + A_{44}\hat{t} = -n^{-1}S_{n4} + o_p(n^{-1/2}). \tag{31}$$

From (31), we have

$$\hat{t} = A_{44}^{-1}\,(A_{41}, A_{42}, A_{43})\,(\hat{\theta}-\theta_0) - A_{44}^{-1}(n^{-1}S_{n4}) + o_p(n^{-1/2}). \tag{32}$$

Substituting (32) into (30) gives

$$\begin{pmatrix} A_{14}A_{44}^{-1}A_{41} & A_{14}A_{44}^{-1}A_{42} & A_{14}A_{44}^{-1}A_{43} \\ A_{24}A_{44}^{-1}A_{41} & A_{24}A_{44}^{-1}A_{42}+A_{22} & A_{24}A_{44}^{-1}A_{43}+A_{23} \\ A_{34}A_{44}^{-1}A_{41} & A_{34}A_{44}^{-1}A_{42}+A_{32} & A_{34}A_{44}^{-1}A_{43}+A_{33} \end{pmatrix} (\hat{\theta}-\theta_0)$$

$$= \begin{pmatrix} 0 & 0 & A_{14}A_{44}^{-1} \\ 1 & 0 & A_{24}A_{44}^{-1} \\ 0 & I_3 & A_{34}A_{44}^{-1} \end{pmatrix} \left(n^{-1}S_n\right) + o_p(n^{-1/2}). \tag{33}$$

Here I_3 denotes the 3×3 identity matrix.

Recall that

$$U = \begin{pmatrix} 0 & 0 & A_{14} \\ A_{22} & A_{23} & A_{24} \\ A_{32} & A_{33} & A_{34} \end{pmatrix} \text{ and } V = \begin{pmatrix} A_{22} & A_{23} & 0 \\ A_{32} & A_{33} & 0 \\ 0 & 0 & A_{44} \end{pmatrix}. \tag{34}$$

After some algebra, it can be verified that the coefficient matrices for $\hat{\theta}-\theta_0$ and $n^{-1}S_n$ in (33) are $J = UV^{-1}U^{\mathrm{T}}$ and UV^{-1}, respectively. Then (33) simplifies to

$$J(\hat{\theta}-\theta_0) = UV^{-1}\Big(n^{-1}S_n\Big) + o_p(n^{-1/2}). \tag{35}$$

Therefore,

$$n^{1/2}(\hat{\theta}-\theta_0) = J^{-1}UV^{-1}\Big(n^{-1/2}S_n\Big) + o_p(1).$$

Via Lemma 3 and the central limit theorem, we conclude that

$$n^{1/2}(\hat{\theta}-\theta_0) \to N\Big(0, \Sigma\Big)$$

in distribution with

$$\Sigma = \{J^{-1}UV^{-1}\}\{\mathrm{var}(n^{-1/2}S_n)\}\{V^{-1}U^{\mathrm{T}}J^{-1}\}. \tag{36}$$

In the last step, we simplify the form of Σ. Let

$$c = \begin{pmatrix} A_{22} \\ A_{32} \\ A_{42}+A_{41} \end{pmatrix}.$$

Then $\mathrm{var}(n^{-1/2}S_n)$ provided in Lemma 3 can be written in the following form:

$$\mathrm{var}(n^{-1/2}S_n) = V - \frac{(1+\rho)^2}{\rho}cc^{\mathrm{T}}. \tag{37}$$

Together with (36), (37) implies that

$$\Sigma = J^{-1} - \frac{(1+\rho)^2}{\rho} J^{-1} U V^{-1} c c^{\mathrm{T}} V^{-1} U^{\mathrm{T}} J^{-1}.$$

Further, note that

$$UV^{-1}c = \begin{pmatrix} 0 & 0 & A_{14}A_{44}^{-1} \\ 1 & 0 & A_{24}A_{44}^{-1} \\ 0 & I_3 & A_{34}A_{44}^{-1} \end{pmatrix} \begin{pmatrix} A_{22} \\ A_{32} \\ A_{42}+A_{41} \end{pmatrix} = \begin{pmatrix} A_{14}A_{44}^{-1}A_{41} + A_{14}A_{44}^{-1}A_{42} \\ A_{24}A_{44}^{-1}A_{41} + A_{24}A_{44}^{-1}A_{42} + A_{22} \\ A_{34}A_{44}^{-1}A_{41} + A_{34}A_{44}^{-1}A_{42} + A_{32} \end{pmatrix},$$

which is the sum of the first two columns of J. Hence,

$$J^{-1}UV^{-1}c = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore,

$$\Sigma = J^{-1} - \frac{(1+\rho)^2}{\rho}(1,1,0,0,0)^{\mathrm{T}}(1,1,0,0,0)$$

as claimed in Theorem 1.

4. PROOF OF COROLLARY 1

Part (a). When $I = 1$ or there is only one estimating equation for the auxiliary information, U becomes a square matrix. Then (35) implies that

$$U^{\mathrm{T}}(\hat{\theta} - \theta_0) = n^{-1} S_n + o_p(n^{-1/2}).$$

That is,

$$\begin{pmatrix} 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \\ A_{41} & A_{42} & A_{43} \end{pmatrix} \begin{pmatrix} \hat{\eta} - \eta_0 \\ \hat{\alpha}^* - \alpha_0^* \\ \hat{\theta}_2 - \theta_{02} \end{pmatrix} = n^{-1} \begin{pmatrix} S_{n2} \\ S_{n3} \\ S_{n4} \end{pmatrix} + o_p(n^{-1/2}).$$

Therefore,

$$\begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} \hat{\alpha}^* - \alpha_0^* \\ \hat{\theta}_2 - \theta_{02} \end{pmatrix} = n^{-1} \begin{pmatrix} S_{n2} \\ S_{n3} \end{pmatrix} + o_p(n^{-1/2}).$$

120 With (7), we further get that

$$\begin{pmatrix} \hat{\alpha}^* - \alpha_0^* \\ \hat{\theta}_2 - \theta_{02} \end{pmatrix} = n^{-1} \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial l(\theta_0, 0, \lambda_0)}{\partial \alpha^*} \\ \frac{\partial l(\theta_0, 0, \lambda_0)}{\partial \theta_2} \end{pmatrix} + o_p(n^{-1/2}). \tag{38}$$

Let $(\hat{\alpha}_L^*, \hat{\theta}_{2L}^{\mathrm{T}})^{\mathrm{T}}$ be the maximum likelihood estimator of $(\alpha^*, \theta_2^{\mathrm{T}})^{\mathrm{T}}$ based on the logistic regression model. Qin & Zhang (1997) showed that $(\hat{\alpha}_L^* - \alpha_0^*, \hat{\theta}_{2L}^{\mathrm{T}} - \theta_{02}^{\mathrm{T}})^{\mathrm{T}}$ has the same approximation as in (38):

$$\begin{pmatrix} \hat{\alpha}_L^* - \alpha_0^* \\ \hat{\theta}_{2L} - \theta_{02} \end{pmatrix} = n^{-1} \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial l(\theta_0, 0, \lambda_0)}{\partial \alpha^*} \\ \frac{\partial l(\theta_0, 0, \lambda_0)}{\partial \theta_2} \end{pmatrix} + o_p(n^{-1/2}).$$

Hence, the asymptotic variances of $n^{1/2}(\hat{\theta}_2 - \theta_{02})$ and $n^{1/2}(\hat{\theta}_{2L} - \theta_{02})$ are the same.

Parts (b, c). For $1 < r \leq I$, let U_r, V_r, J_r, Σ_r denote the corresponding U, V, J, and Σ matrices obtained by using only the first r estimating equations of $g(X, Y; \theta)$. With the result in Part (a), to finish the proof of Parts (b) and (c), it suffices to show that

$$\Sigma_r \leq \Sigma_{r-1}$$

or equivalently

$$J_r \geq J_{r-1}. \tag{39}$$

From (34), we notice that U_r has one more column than U_{r-1}. Let this column be u_r. Then $U_r = (U_{r-1}, u_r)$. Further, using the arguments in the proof of Corollary 1 of Qin & Lawless (1994; p. 318), we get

$$V_r^{-1} \geq \begin{pmatrix} V_{r-1}^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Therefore,

$$J_r = U_r V_r^{-1} U_r^{\tau} \geq (U_{r-1}, u_r) \begin{pmatrix} V_{r-1}^{-1} & 0 \\ 0 & 0 \end{pmatrix} (U_{r-1}, u_r)^{\tau} = U_{r-1} V_{r-1}^{-1} U_{r-1}^{\tau} = J_{r-1},$$

as required by (39). This completes the proof of Corollary 1.

5. Proof of Theorem 2

Recall that $\tilde{\theta}$ is the maximum empirical likelihood estimator of θ under the null hypothesis of $\xi = 0$, and $\tilde{\lambda}$ and $\tilde{t}$ are the corresponding Lagrange multipliers. Note that $\tilde{\theta}$, $\tilde{\lambda}$, and $\tilde{t}$ must satisfy

$$\frac{\partial l(\tilde{\theta}_0, \tilde{t}, \tilde{\lambda})}{\partial \eta} = 0, \quad \frac{\partial l(\tilde{\theta}_0, \tilde{t}, \tilde{\lambda})}{\partial \alpha^*} = 0, \quad \frac{\partial l(\tilde{\theta}_0, \tilde{t}, \tilde{\lambda})}{\partial t} = 0.$$

Via the method used to derive $\hat{\lambda}$ in (6), we similarly have

$$\tilde{\lambda} = \hat{\lambda} = \frac{n_1}{n} = \lambda_0.$$

Hence, in the following derivation we set λ to λ_0.

We now investigate the approximation of $l(\theta, t, \lambda_0)$ when θ and t are in $n^{-1/2}$ neighbourhoods of θ_0 and 0, respectively, which will help us to find an asymptotic approximation of the profile likelihood $\ell(\theta)$.

Applying a second-order Taylor expansion to $l(\theta, t, \lambda_0)$ gives

$$\begin{aligned} l(\theta, t, \lambda_0) = {} & l(\theta_0, 0, \lambda_0) + (\omega - \omega_0)^{\mathrm{T}} \frac{\partial l(\theta_0, 0, \lambda_0)}{\partial \omega} \\ & + \frac{1}{2}(\omega - \omega_0)^{\mathrm{T}} E\left\{ \frac{\partial^2 l(\theta_0, 0, \lambda_0)}{\partial \omega \omega^{\mathrm{T}}} \right\} (\omega - \omega_0) + o_p(1), \end{aligned} \tag{40}$$

where we have used the result in (29) to simplify the second derivatives. Setting the derivative of $l(\theta, t, \lambda_0)$ with respect to t equal to zero and using a similar technique to derive (32), we get

$$t = A_{44}^{-1} (A_{41}, A_{42}, A_{43}) (\theta - \theta_0) - A_{44}^{-1}(n^{-1} S_{n4}) + o_p(n^{-1/2}). \tag{41}$$

Substituting (41) into (40), we get an approximation of the profile likelihood,

$$\ell(\theta) = l(\theta_0, 0, \lambda_0) + (\theta - \theta_0)^{\mathrm{T}} U V^{-1} S_n - \frac{n}{2}(\theta - \theta_0)^{\mathrm{T}} J (\theta - \theta_0) + o_p(1). \tag{42}$$

To investigate the limiting distribution of the empirical likelihood ratio test for testing $H_0 : \xi = 0$, we further profile out $\theta^* = (\eta, \alpha^*, \beta, \gamma)^{\mathrm{T}}$ and obtain the profile likelihood of ξ only. Denote the true value of θ^* by θ_0^*. Let

$$v_n = U V^{-1} S_n = \begin{pmatrix} v_{n1} \\ v_{n2} \end{pmatrix} \tag{43}$$

with v_{n1} consisting of the first four elements of v_n and v_{n2} being the last element of v_n. We partition J accordingly as

$$J = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}.$$

Then the profile likelihood in (42) becomes

$$\begin{aligned} \ell(\theta) = l(\theta_0, 0, \lambda_0) + (\theta^* - \theta_0^*)^{\mathrm{T}} v_{n1} + (\xi - \xi_0) v_{n2} - \frac{n}{2}(\theta^* - \theta_0^*)^{\mathrm{T}} J_{11} (\theta^* - \theta_0^*) \\ - n(\theta^* - \theta_0^*)^{\mathrm{T}} J_{12} (\xi - \xi_0) - \frac{n}{2} J_{22} (\xi - \xi_0)^2 + o_p(1). \end{aligned} \tag{44}$$

Setting the derivative of $\ell(\theta)$ with respect to θ^* equal to zero, we get

$$\theta^* - \theta_0^* = n^{-1} J_{11}^{-1} v_{n1} - J_{11}^{-1} J_{12} (\xi - \xi_0) + o_p(n^{-1/2}). \tag{45}$$

Substituting (45) into (44), we obtain an approximation of the profile likelihood of ξ,

$$\begin{aligned} \ell^*(\xi) = l(\theta_0, 0, \lambda_0) + \frac{n}{2} v_{n1}^{\mathrm{T}} J_{11}^{-1} v_{n1} + (\xi - \xi_0)(v_{n2} - J_{21} J_{11}^{-1} v_{n1}) \\ - \frac{n}{2}(\xi - \xi_0)^2 (J_{22} - J_{21} J_{11}^{-1} J_{12}) + o_p(1). \end{aligned}$$

Then the empirical likelihood ratio test for testing $H_0 : \xi = \xi_0 = 0$ is

$$R(0) = 2\left\{ \sup_{\xi} \ell^*(\xi) - \ell^*(0) \right\} = \frac{(v_{n2} - J_{21} J_{11}^{-1} v_{n1})^2}{n(J_{22} - J_{21} J_{11}^{-1} J_{12})} + o_p(1).$$

To show that the limiting distribution of $R(0)$ is χ_1^2, we need to argue that

$$\frac{v_{n2} - J_{21} J_{11}^{-1} v_{n1}}{\{n(J_{22} - J_{21} J_{11}^{-1} J_{12})\}^{1/2}} \to N(0, 1) \tag{46}$$

in distribution. Using (35) and (43), we note that

$$n^{-1/2}(v_{n2} - J_{21} J_{11}^{-1} v_{n1}) = (-J_{21} J_{11}^{-1}, 1) U V^{-1} (n^{-1/2} S_n) = (-J_{21} J_{11}^{-1}, 1) J (\hat{\theta} - \theta_0) + o_p(1).$$

By Theorem 1, we have

$$n^{-1/2}(v_{n2} - J_{21} J_{11}^{-1} v_{n1}) \to N\Big(0, (-J_{21} J_{11}^{-1}, 1) J \Sigma J (-J_{21} J_{11}^{-1}, 1)^{\mathrm{T}}\Big)$$

in distribution. Therefore, to prove that (46) is correct, we need to verify that

$$(-J_{21} J_{11}^{-1}, 1) J \Sigma J (-J_{21} J_{11}^{-1}, 1)^{\mathrm{T}} = J_{22} - J_{21} J_{11}^{-1} J_{12}. \tag{47}$$

Recall the form of Σ in Theorem 1. We have

$$\begin{aligned}&(-J_{21}J_{11}^{-1},1)J\Sigma J(-J_{21}J_{11}^{-1},1)^{\mathrm{T}}\\ &=(-J_{21}J_{11}^{-1},1)J(-J_{21}J_{11}^{-1},1)^{\mathrm{T}}-\frac{(1+\rho)^2}{\rho}\{(-J_{21}J_{11}^{-1},1)J(1,1,0,0,0)^{\mathrm{T}}\}^2. \qquad (48)\end{aligned}$$

Note that

$$(-J_{21}J_{11}^{-1},1)J=(0,J_{22}-J_{21}J_{11}^{-1}J_{12}).$$

Hence, for the two terms in (48), we have

$$(-J_{21}J_{11}^{-1},1)J(-J_{21}J_{11}^{-1},1)^{\mathrm{T}}=(0,J_{22}-J_{21}J_{11}^{-1}J_{12})(-J_{21}J_{11}^{-1},1)^{\mathrm{T}}=J_{22}-J_{21}J_{11}^{-1}J_{12} \qquad (49)$$

and

$$\{(-J_{21}J_{11}^{-1},1)J(1,1,0,0,0)^{\mathrm{T}}\}^2=\{(0,J_{22}-J_{21}J_{11}^{-1}J_{12})(1,1,0,0,0)^{\mathrm{T}}\}^2=0. \qquad (50)$$

Combining (48)–(50), we verify that (47) is correct. This completes the proof of Theorem 2.

6. MORE SIMULATION RESULTS

6·1. *Sensitivity with biased auxiliary information*

We assess the sensitivity of the proposed test $R(0)$ through simulations under the null $H_0: \xi=0$ using misspecified disease prevalence. Suppose $\tilde{\phi}(a_{i-1},a_i)=\kappa\phi(a_{i-1},a_i)$, $i=1,\ldots,I$, instead of the true prevalence, is used in the test. We consider $\kappa=0{\cdot}90, 0{\cdot}95, 1{\cdot}05$, or $1{\cdot}10$. The other parameters are the same as those used in scenario 1 in the main article. For each κ, the type I error of $R(0)$ at the nominal level 0·05 is estimated using 1000 replications of the simulation, each with 2000 cases and 2000 controls. The results are summarized in Table 1 below, suggesting that our test cannot control the type I error properly if the disease prevalence is incorrectly specified. This is expected; it is equivalent to the model misspecification problem in full parametric inference.

Table 1. *Type* I *error (%) of* $R(0)$ *when we use biased prevalence* $\tilde{\phi}(a_{i-1},a_i)=\kappa\phi(a_{i-1},a_i)$, $i=1,\ldots,I$ *rather than the true prevalence* $\phi(a_{i-1},a_i)$, $i=1,\ldots,I$

κ	0·90	0·95	1·05	1·10
Type I error	8·4	5·9	7· 3	11·4

6·2. *Simulation in the setting of genetic association*

We evaluate the performance of our method in the setting of genetic association. Let X be the continuous environmental risk factor with auxiliary information, and Y the genotype at a binary genetic marker. We code Y as 0, 1, or 2, representing the number of minor alleles in the genotype. We set the minor allele frequency to 0·4. We assume the following risk model:

$$\log\left\{\frac{\mathrm{pr}(D=1\mid x,y)}{\mathrm{pr}(D=0\mid x,y)}\right\}=\alpha+\beta x+\gamma y+\xi xy.$$

In the simulation, we generate X from $N(0,1)$. The main effects of X and Y are set to $\beta = 1{\cdot}00$ and $\gamma = 0{\cdot}08$, respectively. Note that the odds ratio for having one more copy of the minor allele is about 1·08, which is typical in genetic studies of complex diseases. The effect of interaction is set to $\xi = 0{\cdot}08$ under the alternative. The disease prevalence is assumed to be known for X in four intervals $(-\infty, -0{\cdot}67]$, $(-0{\cdot}67, 0]$, $(0, 0{\cdot}67]$, and $(0{\cdot}67, +\infty)$. In this setting, the interaction effect explains about 0·1% of the variance of the disease status (Nagelkerke, 1991), while the main effect of Y and the interaction effect together explain about 0·3%. The chosen level of variation explained by individual genetic markers is very typical in the genetic study of complex trait, such as various common cancers. The odds ratio for having one more copy of higher risk allele is rarely over 1·5 in cancers. It is commonly accepted that risks for most complex diseases are jointly affected by many genetic loci, with each having a very minimal effect. For example, for breast cancer and prostate cancer, the averaged variation explained by a single genetic marker is about 0·25% (Park et al., 2010). We expect that the interaction effect is even weaker, as there are very few interactions, either gene by gene, or gene by environment, that have been detected so far.

We consider the following three scenarios: (1) the true prevalence is known without uncertainty, called scenario S1; (2) the prevalence is estimated through a cohort study from which the case-control study is sampled, called scenario S2; (3) the prevalence is estimated from a separate cohort study, called scenario S3. We estimated the scale parameter in the scaled χ^2_1 distribution using the proposed bootstrap procedure with $B = 500$ in scenarios S2 and S3. The type I errors and powers of our test and the likelihood ratio test based on the logistic regression model are summarized in Table 2. In Table 3, we provide the estimated bias and standard deviation for each of the scenarios. The simulation results suggest that our test can maintain the type I error properly and is more powerful than the likelihood ratio test based on the logistic regression model. Further, the use of auxiliary information can improve the estimated efficiency of β and ξ, but the improvement of γ is limited.

Table 2. *Type* I *error* (%) *and power* (%) *comparison in three scenarios in the setting of genetic association: The results for each scenario are based on* 1000 *simulated datasets, each consisting of* 2000 *cases and* 2000 *controls, with* $(\beta, \gamma) = (1{\cdot}00, 0{\cdot}08)$

	Type I error (%) $\xi = 0$		Power (%) $\xi = 0{\cdot}08$	
Scenario	Logistic	Proposed	Logistic	Proposed
S1	4·7	5·1	26·6	83·5
S2	5·7	5·4	28·7	76·7
S3	5·6	5·5	26·1	77·1

REFERENCES

NAGELKERKE, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika* **78**, 691–692.

PARK, J. H., WACHOLDER, S., GAIL, M. H., PETERS, U., JACOBS, K. B., CHANOCK, S. J. & CHATTERJEE, N. (2010). Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat. Genet.* **42**, 570–575.

Table 3. *Bias and standard deviation (in parentheses) comparison in three scenarios in the setting of genetic association: The results for each scenario are* $1000\times$ *the actual values and are based on* 1000 *simulated datasets, each consisting of* 2000 *cases and* 2000 *controls*

Scenario		$\beta = 1{\cdot}00$	$\gamma = 0{\cdot}08$	$\xi = 0{\cdot}08$
S1	Logistic	0 (70)	-1 (64)	1 (60)
	Proposed	4 (28)	0 (60)	-4 (25)
S2	Logistic	0 (74)	1 (66)	1 (59)
	Proposed	0 (33)	3 (63)	-1 (29)
S3	Logistic	8 (72)	3 (65)	-4 (59)
	Proposed	1 (32)	4 (62)	-2 (29)

QIN, J. & LAWLESS, J. F. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22**, 300–325.
QIN, J. & ZHANG, B. (1997). A goodness-of-fit test for logistic regression models based on case-control data. *Biometrika* **84**, 609–618.