Asymptotic Efficiency of an Exponential Cure Model When Cure Information Is Partially Known

Yu Wu[1], Yong Lin[2], Chin-Shang Li[3], Shou-En Lu[2] & Weichung Joe Shih[2]

[1] K & L Consulting Services, Inc., Fort Washington, PA, USA

[2] Department of Biostatistics, School of Public Health & Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, 683 Hoes Lane West, Piscataway, NJ, USA

[3] Department of Public Health Sciences, Division of Biostatistics, University of California, Davis, CA, USA

Correspondence: Yong Lin, Department of Biostatistics, School of Public Health & Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, 683 Hoes Lane West, Piscataway, NJ 08854, USA. E-mail: linyo@rutgers.edu

Abstract

Cure models are popularly used to analyze failure time data where some individuals could eventually experience and others might never experience an event of interest. However in many studies, there are diagnostic procedures available to provide further information about whether a subject is cured. Wu et al. (2014) proposed a method, called the *extended* cure model, that incorporated such additional diagnostic cured status information into the classical cure model analysis. Through extensive simulations, they demonstrated that the extended cure models provide more efficient and less biased estimations, and higher efficiency and smaller bias are associated with higher sensitivity and specificity of the diagnostic procedure used. In this paper, we provide theoretical justifications of this positive association for some special cases. More specifically we shows that the maximum likelihood estimators (MLEs) of the parameters for an extended exponential cure model are asymptotically more efficient than the MLEs for the corresponding classical exponential cure model.

Keywords: cure model, sensitivity and specificity, asymptotic efficiency

1. Introduction

When there is evidence of long-term survivors, cure models are often used to model the survival curve. Let T be a non-negative random variable for the failure time, $\mathbf{x}$ and $\mathbf{z}$ the covariate vectors, $\pi(\mathbf{z})$ the uncured probability for a subject, and $f(t|\mathbf{x}, \mathbf{z})$ and $S(t|\mathbf{x}, \mathbf{z})$ the probability density function (pdf) and the survival function for T, respectively. Denote $f_u(t|\mathbf{x})$ and $S_u(t|\mathbf{x})$ as the pdf and the survival function for uncured subjects, respectively. The cure model can be written as a mixture model in terms of the pdf: $f(t|\mathbf{x}, \mathbf{z}) = \pi(\mathbf{z})f_u(t|\mathbf{x})$, or in terms of the survival function:

$$S(t|\mathbf{x}, \mathbf{z}) = \pi(\mathbf{z})S_u(t|\mathbf{x}) + [1 - \pi(\mathbf{z})]. \tag{1}$$

In the literature, the cure models have been extensively studied. Conventionally $\pi(\mathbf{z})$ is called the "incidence" part, and $f_u(t|\mathbf{x})$ is referred to as the "latency" part. Logistic regression is commonly used to model the "incidence" part, although other links or non-linear regression methods could be used. The "latency" part can be modeled parametrically, semi-parametrically, or non-parametrically. In the parametric approach, the following distributions have been commonly used: Exponential (Jones et al., 1981; Goldman, 1984; Ghitany & Maller, 1992); Weibull (Farewell, 1982, 1986); Lognormal (Boag, 1949; Gamel et al., 1990); Gompertz (Gordon, 1990a, 1990b; Cantor & Shuster, 1992); Extended generalized gamma (EGG) (Yamaguchi, 1992); and Generalized F (GF) distributions (Peng et al., 1998). In the non-parametric approach, Kaplan-Meier estimation method is used without adjusting for covariates as in Taylor (1995). In the semi-parametric approach, some authors used the Cox proportional hazards (PH) model (Kuk & Chen, 1992; Peng & Dear, 2000; Sy & Taylor, 2000), and some used accelerated failure time (AFT) models (Li & Taylor, 2002; Zhang & Peng, 2007). In general, parametric cure models can achieve greatest efficiency in estimation if the distributional assumptions are satisfied. However in practice it can be challenging to verify these assumptions. Although semi-parametric models do not require a distributional assumption, they may

lose efficiency in estimation compared to a parametric model when a distribution can be correctly identified.

All the cure modeling to date assumes that cured and uncured subjects can not be distinguished in the censored subset. However medical diagnostic procedures in many studies are available to provide further information about whether a subject is cured. For instance, closure of the growth plate can be served as an indicator of cure in the study of bone injury in pediatric patients (Leary et al., 2009; Wu et al., 2014). The diagnostic procedures are likely associated with a certain degree of accuracy in terms of sensitivity and specificity, because it can be difficult to completely separate cured and uncured subjects in the censored subset. Motivated by a clinical study, Wu et al. (2014) extended the classical cure models to incorporate the additional diagnostic information about cured status. Through extensive simulations, they demonstrated that the extended cure models provide more efficient and less biased estimations, and the higher efficiency and smaller bias is associated with higher sensitivity and specificity of diagnostic procedures.

In this paper, we provide theoretical justifications to show how such additional diagnostic information can improve the asymptotic efficiency of model parameter estimators, as compared to the classical cure model approach. Specifically, we provide theoretical justification of this positive association between the sensitivity and specificity of the diagnostic procedure and the asymptotic efficiency of the maximum likelihood estimators (MLEs) of the extended exponential cure model of Wu et al. (2014) in a few special cases.

In Section 2, the formulation of a cure model incorporated with additional cure information (called *extended* cure model) is provided. In Section 3, the asymptotic efficiency of the MLEs of the parameters for an extended exponential cure model and the asymptotic relative efficiency (ARE) of the MLEs respect to the MLEs for the traditional exponential cure model are systematically studied under some special cases. Discussion is given in Section 4.

2. Extended Cure Models

Extended cure models have been introduced by Wu et al. (2014). Let $O_1 = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i), i = 1, 2, \ldots, n\}$ be a data set. Here t_i is the observed survival time of subject i, δ_i is the censoring indicator with 1 if t_i is uncensored (i.e., observed), and 0 otherwise, $\mathbf{x}_i$ and $\mathbf{z}_i$ are two covariate vectors. Let $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ be the parameter vectors related to $\mathbf{x}_i$ and $\mathbf{z}_i$, respectively, and $\boldsymbol{\theta}'_1 = (\boldsymbol{\beta}', \boldsymbol{\gamma}')$. If the cure model in (1) is used for modeling the data set O_1, the observed likelihood can be written as:

$$L_o(\boldsymbol{\theta}_1; O_1) = \prod_{i=1}^{n} [\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i)]^{\delta_i} \{\pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i) + [1 - \pi(\mathbf{z}_i)]\}^{1-\delta_i}. \tag{2}$$

Assume that for censored subjects, their diagnostic results d_i are also observed, where d_i is 1 if subject i is diagnosed as cured and 0 if diagnosed as uncured. A diagnostic procedure usually is associated with certain sensitivity and specificity. Sensitivity measures the proportion of actual positives which are correctly identified (e.g., the percentage of sick people who are correctly identified as sick). Specificity measures the proportion of actual negatives who are correctly identified (e.g., the percentage of healthy people who are correctly identified as healthy). Suppose that the diagnostic procedure results are independent of the failure times, i.e., d_i is independent of t_i, and the diagnostic procedure has a sensitivity of p_0 and a specificity of $1 - p_1$. We have $p_0 \geq p_1$ for a validated diagnostic procedure. Although p_0 and p_1 might be modeled, for simplicity they are assumed not to depend on any covariates. Let $O_2 = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i, d_i), i = 1, 2, \ldots, n\}$ and $\boldsymbol{\theta}'_2 = (\boldsymbol{\theta}'_1, p_0, p_1)$. For uncensored individuals ($\delta_i = 1$), the contribution to the likelihood is the same as that in (2); while for censored individuals ($\delta_i = 0$), with the independent assumption of d_i and t_i, the contribution is $p_1^{d_i}(1-p_1)^{1-d_i}\pi(\mathbf{z}_i)S_u(t_i|\mathbf{x}_i)$ if they are uncured, and the contribution is $p_0^{d_i}(1-p_0)^{1-d_i}[1-\pi(\mathbf{z}_i)]$ if they are cured. A cure model incorporated with these additional diagnostic information will be called an *extended* cure model. The observed likelihood for the extended cure model is as follows:

$$L_o(\boldsymbol{\theta}_2; O_2) = \prod_{i=1}^{n} [\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i)]^{\delta_i} \left\{ p_1^{d_i}(1-p_1)^{1-d_i}\pi(\mathbf{z}_i)S_u(t_i|\mathbf{x}_i) + p_0^{d_i}(1-p_0)^{1-d_i}[1-\pi(\mathbf{z}_i)] \right\}^{1-\delta_i}. \tag{3}$$

Because the diagnostic procedure results may not always be available for all the censored subjects, let $\eta_i = 1$ if the diagnostic result of subject i is available, and $\eta_i = 0$ otherwise. Let $O_3 = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i, \eta_i, d_i), i = 1, 2, \ldots, n\}$. We can then write the observed likelihood for the extended cure model when cure information is partially known as follows:

$$L_o(\boldsymbol{\theta}_2; \boldsymbol{O}_3) = \prod_{i=1}^{n} [\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i)]^{\delta_i} \times \{p_1^{d_i}(1-p_1)^{1-d_i}\pi(\mathbf{z}_i)S_u(t_i|\mathbf{x}_i) + p_0^{d_i}(1-p_0)^{1-d_i}[1-\pi(\mathbf{z}_i)]\}^{(1-\delta_i)\eta_i}$$

$$\{\pi(\mathbf{z}_i)S_u(t_i|\mathbf{x}_i) + [1-\pi(\mathbf{z}_i)]\}^{(1-\delta_i)(1-\eta_i)}. \tag{4}$$

It is noted that (4) reduces to (2) except for a constant multiplier when $p_0 = p_1$, which means that if both sensitivity and $(1 - \text{specificity})$ are the same, the likelihood functions with and without the diagnostic information are the same. In practice, we want both sensitivity and specificity to be high and $p_0 \neq p_1$.

As in the literature, one can use logistic regression, other link functions or nonlinear regression to model the "incidence" part $\pi(\mathbf{z})$. Parametric, semiparametric (PH or AFT), or nonparametric methods can be used to model the "latency" part $S_u(t|\mathbf{x})$. An expectation-maximization (EM) algorithm can be used to estimate the model parameters in (4). The details of the EM procedure can be found in Wu et al. (2014). In this paper, we focus on the asymptotic efficiency of the MLEs of the parameters in the extended exponential cure model with the observed likelihood in Equation (3).

3. Asymptotic Efficiency of Maximum Likelihood Estimation for Extended Exponential Cure Models

In this section, we show for several special cases that the asymptotic efficiencies of the MLEs for an extended exponential cure model are positively associated with the sensitivity and the specificity of the diagnostic procedure, and are asymptotically more efficient than the MLEs for the corresponding classical cure model. Assume that the logit link is used for the incidence part, the exponential distribution for the latency part, and p_0 and p_1 are known. Specifically, the assumptions are stated as follows:

- $\log\left(\frac{\pi(\mathbf{z}_i)}{1-\pi(\mathbf{z}_i)}\right) = \boldsymbol{\gamma}'\mathbf{z}_i$, where $\boldsymbol{\gamma}' = (\gamma_0, \gamma_1, \ldots, \gamma_k)$ is a $1 \times (k+1)$ parameter vector, and $\mathbf{z}_i = (z_{i0}, z_{i1}, \ldots, z_{ik})'$ is a $(k+1) \times 1$ covariate vector with $z_{i0} = 1$.

- $f_u(t_i|\mathbf{x}_i) = h(\mathbf{x}_i)e^{-h(\mathbf{x}_i)t_i}$ is the pdf of an exponential distribution and $h(\mathbf{x}_i) = e^{\boldsymbol{\beta}'\mathbf{x}_i}$. Here $\mathbf{x}_i = (x_{i0}, x_{i1}, \ldots, x_{im})'$ is a $(m+1) \times 1$ covariate vector with $x_{i0} = 1$. $\boldsymbol{\beta}' = (\beta_0, \beta_1, \ldots, \beta_m)$ is a $1 \times (m+1)$ parameter vector.

- p_0 and p_1 are known with $p_0 \geq p_1$ for a valid diagnostic procedure.

Proposition 1 *Denote $V_{\boldsymbol{\gamma}}^D$ as the asymptotic variance of the MLE of $\boldsymbol{\gamma}$ when the diagnostic procedure is used, and $V_{\boldsymbol{\gamma}}^N$ as the asymptotic variance of the MLE of $\boldsymbol{\gamma}$ when no diagnostic procedure is used. Let $V_{\boldsymbol{\beta}}^D$ be the asymptotic variance of the MLE of $\boldsymbol{\beta}$ when the diagnostic procedure is used, and $V_{\boldsymbol{\beta}}^N$ the asymptotic variance of the MLE of $\boldsymbol{\beta}$ when no diagnostic procedure is used. The following results are true:*

(1) When sensitivity and specificity are both 100%, i.e., $p_0 = 1$, $p_1 = 0$, all diagonal entries of $V_{\boldsymbol{\gamma}}^D$ and $V_{\boldsymbol{\beta}}^D$ are less than or equal to the corresponding entries of $V_{\boldsymbol{\gamma}}^N$ and $V_{\boldsymbol{\beta}}^N$. This implies that the estimators of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are more efficient when diagnostic information is included.

(2) When $k = 0$, $m = 0$, i.e., $\boldsymbol{\gamma} = (\gamma_0)$, $\boldsymbol{\beta} = (\beta_0)$, $V_{\boldsymbol{\gamma}}^D$ and $V_{\boldsymbol{\beta}}^D$ are less than or equal to $V_{\boldsymbol{\gamma}}^N$ and $V_{\boldsymbol{\beta}}^N$, respectively. This implies that the estimators of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are more efficient when diagnostic information is included. Furthermore, the asymptotic variance decreases as the sensitivity or specificity increases.

(3) When $k = 0$, $m = 1$, i.e., $\boldsymbol{\gamma} = (\gamma_0)$, $\boldsymbol{\beta} = (\beta_0, \beta_1)'$, and x_{i1} is a binary variable with values of 0 and 1, the asymptotic variances of the MLEs of γ_0 and β_0 are smaller when the diagnostic procedure is used. This implies that the estimators of γ_0 and β_0 are more efficient when diagnostic information is included. Furthermore, the asymptotic variance decreases as the sensitivity or specificity increases.

(4) When $k = 1$, $m = 0$, i.e., $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)'$, $\boldsymbol{\beta} = (\beta_0)$, and z_{i1} is a binary variable with values of 0 and 1, the asymptotic variances of the MLEs of γ_0 and β_0 are smaller when the diagnostic procedure is used. This implies that the estimators of γ_0 and β_0 are more efficient when diagnostic information is included. Furthermore, the asymptotic variance decreases as the sensitivity or specificity increases.

The proposition will be proved based on several Lemmas. For convenience, for all the derivations in this section, denote $\pi_i = \pi(\mathbf{z}_i)$ and $h_i = h(\mathbf{x}_i)$. The observed likelihood for the extended exponential cure model according to (3) can be written as:

$$L_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2) = \prod_{i=1}^{n} (\pi_i h_i e^{-h_i t_i})^{\delta_i} [p_1^{d_i}(1-p_1)^{1-d_i}\pi_i e^{-h_i t_i} + p_0^{d_i}(1-p_0)^{1-d_i}(1-\pi_i)]^{1-\delta_i}, \tag{5}$$

which implies that the observed log-likelihood is:

$$
\ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2) = \log[L_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)]
$$

$$
= \sum_{i=1}^{n} \delta_i[\log(\pi_i) + \log(h_i) - h_i t_i] + \sum_{i=1}^{n} (1 - \delta_i) \log[p_1^{d_i}(1 - p_1)^{1-d_i}\pi_i e^{-h_i t_i} + p_0^{d_i}(1 - p_0)^{1-d_i}(1 - \pi_i)]. \quad (6)
$$

The score functions are:

$$
\frac{\partial \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \boldsymbol{\gamma}} = \frac{\partial \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \pi_i}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} = \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\left[\frac{\delta_i}{\pi_i} + (1 - \delta_i)\frac{p_1^{d_i}(1 - p_1)^{1-d_i}e^{-h_i t_i} - p_0^{d_i}(1 - p_0)^{1-d_i}}{p_1^{d_i}(1 - p_1)^{1-d_i}\pi_i e^{-h_i t_i} + p_0^{d_i}(1 - p_0)^{1-d_i}(1 - \pi_i)}\right] \quad (7)
$$

and

$$
\frac{\partial \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \boldsymbol{\beta}} = \frac{\partial \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial h_i}\frac{\partial h_i}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}}\left[\frac{\delta_i}{h_i} - \delta_i t_i - (1 - \delta_i)\frac{p_1^{d_i}(1 - p_1)^{1-d_i}\pi_i t_i e^{-h_i t_i}}{p_1^{d_i}(1 - p_1)^{1-d_i}\pi_i e^{-h_i t_i} + p_0^{d_i}(1 - p_0)^{1-d_i}(1 - \pi_i)}\right]. \quad (8)
$$

By defining

$$
a_i = p_0^{d_i}(1 - p_0)^{1-d_i}, \ b_i = p_1^{d_i}(1 - p_1)^{1-d_i}, \ \text{and} \ v_i = \frac{b_i}{a_i},
$$

one can simplify (7) and (8) to

$$
\frac{\partial \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \boldsymbol{\gamma}} = \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\left[\frac{\delta_i}{\pi_i} + (1 - \delta_i)\frac{v_i e^{-h_i t_i} - 1}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i}\right],
$$

$$
\frac{\partial \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}}\left[\frac{\delta_i}{h_i} - \delta_i t_i - (1 - \delta_i)\frac{v_i \pi_i t_i e^{-h_i t_i}}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i}\right].
$$

The entries of the observed information matrix are

$$
\mathbf{I}_{11} = -\frac{\partial^2 \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\gamma}'}
$$

$$
= \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\left[\frac{\delta_i}{\pi_i^2} + (1 - \delta_i)\frac{(v_i e^{-h_i t_i} - 1)^2}{(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right] - \sum_{i=1}^{n} \frac{\partial^2 \pi_i}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\gamma}'}\left[\frac{\delta_i}{\pi_i} + (1 - \delta_i)\frac{v_i e^{-h_i t_i} - 1}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i}\right], \quad (9)
$$

$$
\mathbf{I}_{22} = -\frac{\partial^2 \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}
$$

$$
= \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[\frac{\delta_i}{h_i^2} - (1 - \delta_i)\frac{v_i \pi_i(1 - \pi_i)t_i^2 e^{-h_i t_i}}{(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right] - \sum_{i=1}^{n} \frac{\partial^2 h_i}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}\left[\frac{\delta_i}{h_i} - \delta_i t_i - (1 - \delta_i)\frac{v_i \pi_i t_i e^{-h_i t_i}}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i}\right], \quad (10)
$$

$$
\mathbf{I}_{12} = -\frac{\partial^2 \ell_o(\boldsymbol{\theta}_2; \boldsymbol{O}_2)}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\beta}'} = \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[(1 - \delta_i)\frac{v_i t_i e^{-h_i t_i}}{(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right]. \quad (11)
$$

For any γ_m and γ_n, and for observation i, because $\pi_i = \frac{e^{\gamma' z_i}}{1+e^{\gamma' z_i}}$, the first order partial derivatives of π_i are

$$
\frac{\partial \pi_i}{\partial \gamma_m} = \frac{z_{im}e^{\gamma' \mathbf{z}_i}}{(1 + e^{\gamma' \mathbf{z}_i})^2}, \ \frac{\partial \pi_i}{\partial \gamma_m}\frac{\partial \pi_i}{\partial \gamma_n} = \frac{z_{im}z_{in}e^{2\gamma' \mathbf{z}_i}}{(1 + e^{\gamma' \mathbf{z}_i})^4}, \ \text{and} \ \left(\frac{\partial \pi_i}{\partial \gamma_m}\right)^2 = \frac{z_{im}{}^2 e^{2\gamma' \mathbf{z}_i}}{(1 + e^{\gamma' \mathbf{z}_i})^4}.
$$

The second order partial derivatives of π_i are

$$
\frac{\partial^2 \pi_i}{\partial \gamma_m^2} = \frac{1 - e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}\left(\frac{\partial \pi_i}{\partial \gamma_m}\right)^2 \quad (12)
$$

and

$$
\frac{\partial^2 \pi_i}{\partial \gamma_m \partial \gamma_n} = \frac{1 - e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}\frac{\partial \pi_i}{\partial \gamma_m}\frac{\partial \pi_i}{\partial \gamma_n}. \quad (13)
$$

From (12) and (13), we have

$$\frac{\partial^2 \pi_i}{\partial \gamma \partial \gamma'} = \frac{1 - e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}} \frac{\partial \pi_i}{\partial \gamma} \frac{\partial \pi_i}{\partial \gamma'}.$$

Similarly for any β_m and β_n, and for observation i, the first order partial derivatives of $h_i = e^{\beta' \mathbf{x}_i}$ are

$$\frac{\partial h_i}{\partial \beta_m} = x_{im} e^{\beta' \mathbf{x}_i}, \ \frac{\partial h_i}{\partial \beta_m} \frac{\partial h_i}{\partial \beta_n} = x_{im} x_{in} e^{2\beta' \mathbf{x}_i}, \ \text{and} \ \left(\frac{\partial h_i}{\partial \beta_m} \right)^2 = x_{im}^2 e^{2\beta' \mathbf{x}_i}.$$

The second order partial derivatives of h_i are

$$\frac{\partial^2 h_i}{\partial \beta_m^2} = x_{im}^2 e^{\beta' \mathbf{x}_i} = \frac{1}{h_i} \left(\frac{\partial h_i}{\partial \beta_m} \right)^2, \tag{14}$$

$$\frac{\partial^2 h_i}{\partial \beta_m \partial \beta_n} = x_{im} x_{in} e^{\beta' \mathbf{x}_i} = \frac{1}{h_i} \frac{\partial h_i}{\partial \beta_m} \frac{\partial h_i}{\partial \beta_n}. \tag{15}$$

From (14) and (15), we have

$$\frac{\partial^2 h_i}{\partial \beta \partial \beta'} = \frac{1}{h_i} \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'}.$$

Consequently, $\mathbf{I}_{11}$ in (9) and $\mathbf{I}_{22}$ in (10) can be rewritten as follows:

$$\mathbf{I}_{11} = \sum_{i=1}^n \frac{\partial \pi_i}{\partial \gamma} \frac{\partial \pi_i}{\partial \gamma'} \left[\frac{\delta_i}{\pi_i^2} + (1 - \delta_i) \frac{(v_i e^{-h_i t_i} - 1)^2}{(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i)^2} \right] - \sum_{i=1}^n \frac{\partial \pi_i}{\partial \gamma} \frac{\partial \pi_i}{\partial \gamma'} \frac{1 - e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}} \left[\frac{\delta_i}{\pi_i} + (1 - \delta_i) \frac{v_i e^{-h_i t_i} - 1}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i} \right] \tag{16}$$

and

$$\mathbf{I}_{22} = \sum_{i=1}^n \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'} \left[\frac{\delta_i}{h_i^2} - (1 - \delta_i) \frac{v_i \pi_i (1 - \pi_i) t_i^2 e^{-h_i t_i}}{(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i)^2} \right] - \sum_{i=1}^n \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'} \frac{1}{h_i} \left[\frac{\delta_i}{h_i} - \delta_i t_i - (1 - \delta_i) \frac{v_i \pi_i t_i e^{-h_i t_i}}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i} \right]. \tag{17}$$

Similarly, if no diagnostic information is used, we only need to set $v_i = 1$ or $p_0 = p_1 = 0.5$ in (16), (17), and (11) to have the following entries

$$\mathbf{J}_{11} = -\frac{\partial^2 \ell_o(\theta_1; \mathbf{O}_1)}{\partial \gamma \partial \gamma'}$$
$$= \sum_{i=1}^n \frac{\partial \pi_i}{\partial \gamma} \frac{\partial \pi_i}{\partial \gamma'} \left[\frac{\delta_i}{\pi_i^2} + (1 - \delta_i) \frac{(e^{-h_i t_i} - 1)^2}{(\pi_i e^{-h_i t_i} + 1 - \pi_i)^2} \right] - \sum_{i=1}^n \frac{\partial \pi_i}{\partial \gamma} \frac{\partial \pi_i}{\partial \gamma'} \frac{1 - e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}} \left[\frac{\delta_i}{\pi_i} + (1 - \delta_i) \frac{e^{-h_i t_i} - 1}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right],$$
$$\mathbf{J}_{22} = -\frac{\partial^2 \ell_o(\theta_1; \mathbf{O}_1)}{\partial \beta \partial \beta'}$$
$$= \sum_{i=1}^n \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'} \left[\frac{\delta_i}{h_i^2} - (1 - \delta_i) \frac{\pi_i (1 - \pi_i) t_i^2 e^{-h_i t_i}}{(\pi_i e^{-h_i t_i} + 1 - \pi_i)^2} \right] - \sum_{i=1}^n \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'} \frac{1}{h_i} \left[\frac{\delta_i}{h_i} - \delta_i t_i - (1 - \delta_i) \frac{\pi_i t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right],$$
$$\mathbf{J}_{12} = -\frac{\partial^2 \ell_o(\theta_1; \mathbf{O}_1)}{\partial \gamma \partial \beta'} = \sum_{i=1}^n \frac{\partial \pi_i}{\partial \gamma} \frac{\partial h_i}{\partial \beta'} \left[(1 - \delta_i) \frac{t_i e^{-h_i t_i}}{(\pi_i e^{-h_i t_i} + 1 - \pi_i)^2} \right].$$

Denote $\mathbf{T} = \{t_i, i = 1, 2, \ldots, n\}$ and $\mathbf{V} = \{(\delta_i, d_i), i = 1, 2, \ldots, n\}$. To obtain the information matrix, we will take expectation of $\mathbf{I}_{rs}$ and $\mathbf{J}_{rs}$, $r, s = 1, 2$, with respect to $\mathbf{O} = \{\mathbf{T}, \mathbf{V}\}$. Let

$$\varphi_i(p_0, p_1) = \frac{(p_0 - p_1)^2}{[(1 - p_1) \pi_i e^{-h_i t_i} + (1 - p_0)(1 - \pi_i)][p_1 \pi_i e^{-h_i t_i} + p_0 (1 - \pi_i)]}. \tag{18}$$

We have the following results.

Lemma 2 *Denote* $\mathbf{I}_{12}^{(i)}$ *and* $\mathbf{J}_{12}^{(i)}$ *as the ith summand of* $\mathbf{I}_{12}$ *and* $\mathbf{J}_{12}$, *respectively. Then*

$$\Delta_{12}^{(i)} = E_O\left(\mathbf{I}_{12}^{(i)}\right) - E_O\left(\mathbf{J}_{12}^{(i)}\right) = -\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{t_i\pi_i e^{-2h_i t_i}(1-\pi_i)}{\pi_i e^{-h_i t_i} + 1 - \pi_i}\varphi_i(p_0, p_1)\right\}.$$

Proof. Because for each i

$$v_i = \frac{p_1^{d_i}(1-p_1)^{1-d_i}}{p_0^{d_i}(1-p_0)^{1-d_i}} = \begin{cases} \frac{1-p_1}{1-p_0} & \text{if } d_i = 0 \\ \frac{p_1}{p_0} & \text{if } d_i = 1 \end{cases},$$

$$\begin{aligned}
E_O\left(\mathbf{I}_{12}^{(i)}\right) &= E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[(1-\delta_i)\frac{v_i t_i e^{-h_i t_i}}{(v_i\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right]\right\} \\
&= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[(1-\delta_i)\frac{v_i t_i e^{-h_i t_i}}{(v_i\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right]\right\}\right\} \\
&= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[\frac{(1-\delta_i)(1-d_i)v_i t_i e^{-h_i t_i}}{(v_i\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right]\right\}\right\} + E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[\frac{(1-\delta_i)d_i v_i t_i e^{-h_i t_i}}{(v_i\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right]\right\}\right\} \\
&= E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{(1-\delta_i)(1-d_i)v_i t_i e^{-h_i t_i}}{(v_i\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right]\right\}\right\} + E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{(1-\delta_i)d_i v_i t_i e^{-h_i t_i}}{(v_i\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right]\right\}\right\} \\
&= E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}P(\delta_i = 0, d_i = 0|t_i)\frac{\frac{1-p_1}{1-p_0}t_i e^{-h_i t_i}}{(\frac{1-p_1}{1-p_0}\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right\} \\
&\quad + E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}P(\delta_i = 0, d_i = 1|t_i)\frac{\frac{p_1}{p_0}t_i e^{-h_i t_i}}{(\frac{p_1}{p_0}\pi_i e^{-h_i t_i} + 1 - \pi_i)^2}\right\} \\
&= E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}P(\delta_i = 0, d_i = 0|t_i)\frac{(1-p_1)(1-p_0)t_i e^{-h_i t_i}}{[(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)]^2}\right\} \\
&\quad + E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}P(\delta_i = 0, d_i = 1|t_i)\frac{p_1 p_0 t_i e^{-h_i t_i}}{[p_1\pi_i e^{-h_i t_i} + p_0(1-\pi_i)]^2}\right\},
\end{aligned} \tag{19}$$

$$P(\delta_i = 0, d_i = 0|t_i) = (1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i), \tag{20}$$

and

$$P(\delta_i = 0, d_i = 1|t_i) = p_1\pi_i e^{-h_i t_i} + p_0(1-\pi_i), \tag{21}$$

by plugging (20) and (21) into (19), we have

$$\begin{aligned}
E_O\left(\mathbf{I}_{12}^{(i)}\right) &= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{[(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)]\frac{(1-p_1)(1-p_0)t_i e^{-h_i t_i}}{[(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)]^2}\right\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{[p_1\pi_i e^{-h_i t_i} + p_0(1-\pi_i)] \times \frac{p_1 p_0 t_i e^{-h_i t_i}}{[p_1\pi_i e^{-h_i t_i} + p_0(1-\pi_i)]^2}\right\} \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left[\frac{(1-p_1)(1-p_0)t_i e^{-h_i t_i}}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)}\right] + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left[\frac{p_1 p_0 t_i e^{-h_i t_i}}{p_1\pi_i e^{-h_i t_i} + p_0(1-\pi_i)}\right] \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{t_i e^{-h_i t_i}\left[\frac{(1-p_1)(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)} + \frac{p_1 p_0}{p_1\pi_i e^{-h_i t_i} + p_0(1-\pi_i)}\right]\right\}.
\end{aligned} \tag{22}$$

Similarly, we have

$$E_O\left(\mathbf{J}_{12}^{(i)}\right) = \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i}\right\}. \tag{23}$$

It can be shown from (22) and (23) that

$$
\begin{aligned}
\Delta_{12}^{(i)} &= E_O\left(\mathbf{I}_{12}^{(i)}\right) - E_O\left(\mathbf{J}_{12}^{(i)}\right) \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left\{t_i e^{-h_i t_i}\left[\frac{(1-p_1)(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i}+(1-p_0)(1-\pi_i)}\right.\right. \\
&\qquad\qquad \left.\left. +\frac{p_1 p_0}{p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)}\right]\right\} - \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left[\frac{t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i}+(1-\pi_i)}\right] \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left\{t_i e^{-h_i t_i}\left[\frac{(1-p_1)(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i}+(1-p_0)(1-\pi_i)} -\frac{1-p_0}{\pi_i e^{-h_i t_i}+(1-\pi_i)}\right]\right\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left\{t_i e^{-h_i t_i}\left[\frac{p_1 p_0}{p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)} -\frac{p_0}{\pi_i e^{-h_i t_i}+(1-\pi_i)}\right]\right\} \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left\{\frac{t_i e^{-h_i t_i}}{\pi_i e^{-h_i}t_i+(1-\pi_i)} \times \frac{(1-p_0)(1-\pi_i)(p_0-p_1)}{(1-p_1)\pi_i e^{-h_i t_i}+(1-p_0)(1-\pi_i)}\right\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left\{\frac{t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i}+(1-\pi_i)} \times \frac{p_0(1-\pi_i)(p_1-p_0)}{p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)}\right\} \\
&= -\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left\{\frac{t_i \pi_i e^{-2h_i t_i}(1-\pi_i)}{\pi_i e^{-h_i t_i}+(1-\pi_i)}\frac{(p_0-p_1)^2}{[(1-p_1)\pi_i e^{-h_i t_i}+(1-p_0)(1-\pi_i)][p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)]}\right\} \\
&= -\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta'}} E_\mathbf{T}\left\{\frac{t_i \pi_i e^{-2h_i t_i}(1-\pi_i)}{\pi_i e^{-h_i t_i}+1-\pi_i}\varphi_i(p_0,p_1)\right\}.
\end{aligned}
$$

$\square$

Lemma 3 *Denote* $\mathbf{I}_{11}^{(i)}$ *and* $\mathbf{J}_{11}^{(i)}$ *as the* i^{th} *summand of* $\mathbf{I}_{11}$ *and* $\mathbf{J}_{11}$*, respectively. Then*

$$
\Delta_{11}^{(i)} = E_O\left(\mathbf{I}_{11}^{(i)}\right) - E_O\left(\mathbf{J}_{11}^{(i)}\right) = \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}} E_\mathbf{T}\left\{\frac{e^{-2h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}\varphi_i(p_0,p_1)\right\}.
$$

Proof. First of all, $\Delta_{11}^{(i)}$ can be expressed as follows:

$$
\begin{aligned}
\Delta_{11}^{(i)} &= E_O\left(\mathbf{I}_{11}^{(i)}\right) - E_O\left(\mathbf{J}_{11}^{(i)}\right) \\
&= E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\left[(1-\delta_i)\frac{(v_i e^{-h_i t_i}-1)^2}{(v_i \pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\} - E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\left[(1-\delta_i)\frac{(e^{-h_i t_i}-1)^2}{(\pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\} \\
&\quad - E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\frac{1-e^{2\boldsymbol{\gamma'}\mathbf{z}_i}}{e^{\boldsymbol{\gamma'}\mathbf{z}_i}}\left[(1-\delta_i)\frac{v_i e^{-h_i t_i}-1}{v_i \pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\} + E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\frac{1-e^{2\boldsymbol{\gamma'}\mathbf{z}_i}}{e^{\boldsymbol{\gamma'}\mathbf{z}_i}}\left[(1-\delta_i)\frac{e^{-h_i t_i}-1}{\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\}.
\end{aligned}
$$

We can write the third term in the above expression as follows:

$$
\begin{aligned}
&E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\frac{1-e^{2\boldsymbol{\gamma'}\mathbf{z}_i}}{e^{\boldsymbol{\gamma'}\mathbf{z}_i}}\left[(1-\delta_i)\frac{v_i e^{-h_i t_i}-1}{v_i \pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\} \\
=&E_\mathbf{T}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\frac{1-e^{2\boldsymbol{\gamma'}\mathbf{z}_i}}{e^{\boldsymbol{\gamma'}\mathbf{z}_i}}\left[(1-\delta_i)(1-d_i)\frac{v_i e^{-h_i t_i}-1}{v_i \pi_i e^{-h_i t_i}+(1-\pi_i)}\right]\right\}\right\} \\
&\quad + E_\mathbf{T}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\frac{1-e^{2\boldsymbol{\gamma'}\mathbf{z}_i}}{e^{\boldsymbol{\gamma'}\mathbf{z}_i}}\left[(1-\delta_i)d_i\frac{v_i e^{-h_i t_i}-1}{v_i \pi_i e^{-h_i t_i}+p_0(1-\pi_i)}\right]\right\}\right\} \\
=&E_\mathbf{T}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\frac{1-e^{2\boldsymbol{\gamma'}\mathbf{z}_i}}{e^{\boldsymbol{\gamma'}\mathbf{z}_i}}P(\delta_i=0,d_i=0|t_i)\frac{(1-p_1)e^{-h_i t_i}-(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i}+(1-p_0)(1-\pi_i)}\right\} \\
&\quad + E_\mathbf{T}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma'}}\frac{1-e^{2\boldsymbol{\gamma'}\mathbf{z}_i}}{e^{\boldsymbol{\gamma'}\mathbf{z}_i}}P(\delta_i=0,d_i=1|t_i)\frac{p_1 e^{-h_i t_i}-p_0}{p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)}\right\}.
\end{aligned}
$$

By using the expressions of $P(\delta_i=0,d_i=0|t_i)$ and $P(\delta_i=0,d_i=1|t_i)$ in (20) and (21), respectively, we can

simplify the third term as follows:

$$
E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{1-e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}\left[(1-\delta_i)\frac{v_i e^{-h_i t_i}-1}{v_i \pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\}
$$
$$
=E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{1-e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}[(1-p_1)e^{-h_i t_i}-(1-p_0)]\right\}+E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{1-e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}(p_1 e^{-h_i t_i}-p_0)\right\}
$$
$$
=E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{1-e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}(e^{-h_i t_i}-1)\right\}.
$$

The above expression does not depend on v_i, so it turns out that

$$
E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{1-e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}\left[(1-\delta_i)\frac{v_i e^{-h_i t_i}-1}{v_i \pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\}=E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{1-e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}}\left[(1-\delta_i)\frac{e^{-h_i t_i}-1}{\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\}.
$$

It follows that

$$
\Delta_{11}^{(i)}=E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\left[(1-\delta_i)\frac{(v_i e^{-h_i t_i}-1)^2}{(v_i \pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\}-E_O\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\left[(1-\delta_i)\frac{(e^{-h_i t_i}-1)^2}{(\pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\}
$$
$$
=E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{(1-\delta_i)(1-d_i)(v_i e^{-h_i t_i}-1)^2}{(v_i \pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\}+E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{(1-\delta_i)d_i(v_i e^{-h_i t_i}-1)^2}{(v_i \pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\}
$$
$$
-E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{(1-\delta_i)(e^{-h_i t_i}-1)^2}{(\pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\}
$$
$$
=E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{P(\delta_i=0,d_i=0|t_i)[(1-p_1)e^{-h_i t_i}-(1-p_0)]^2}{[(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)(1-p_0)]^2}\right\}
$$
$$
+E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{P(\delta_i=0,d_i=1|t_i)[p_1 e^{-h_i t_i}-p_0]^2}{[p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)]^2}\right\}-E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\frac{P(\delta_i=0|t_i)(e^{-h_i t_i}-1)^2}{(\pi_i e^{-h_i t_i}+1-\pi_i)^2}\right\}.
$$

Again because of (20) and (21), we have

$$
\Delta_{11}^{(i)}=\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{[(1-p_1)e^{-h_i t_i}-(1-p_0)]^2}{(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)(1-p_0)}\right\}
$$
$$
+\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{(p_1 e^{-h_i t_i}-p_0)^2}{p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)}\right\}-\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{(e^{-h_i t_i}-1)^2}{\pi_i e^{-h_i t_i}+1-\pi_i}\right\}
$$
$$
=\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{[(1-p_1)e^{-h_i t_i}-(1-p_0)]^2}{(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)(1-p_0)}-\frac{(1-p_1)(e^{-h_i t_i}-1)^2}{\pi_i e^{-h_i t_i}+1-\pi_i}\right\}
$$
$$
+\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{(p_1 e^{-h_i t_i}-p_0)^2}{p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)}-\frac{p_1(e^{-h_i t_i}-1)^2}{\pi_i e^{-h_i t_i}+1-\pi_i}\right\}
$$
$$
=\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{p_1-p_0}{\pi_i e^{-h_i t_i}+1-\pi_i}\frac{-(1-p_1)(1+\pi_i)e^{-2h_i t_i}+(2-p_0-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)(1-p_0)}\right\}
$$
$$
+\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{p_0-p_1}{\pi_i e^{-h_i t_i}+1-\pi_i}\frac{-p_1(1+\pi_i)e^{-2h_i t_i}+(p_0+p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)p_0}{p_1 \pi_i e^{-h_i t_i}+p_0(1-\pi_i)}\right\}
$$
$$
=\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}E_{\mathbf{T}}\left\{\frac{e^{-2h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}\varphi_i(p_0,p_1)\right\}.
$$

$\square$

Lemma 4 *Denote* $\mathbf{I}_{22}^{(i)}$ *and* $\mathbf{J}_{22}^{(i)}$ *as the* i^{th} *summand of* $\mathbf{I}_{22}$ *and* $\mathbf{J}_{22}$, *respectively. Then*

$$
\Delta_{22}^{(i)}=E_O\left(\mathbf{I}_{22}^{(i)}\right)-E_O\left(\mathbf{J}_{22}^{(i)}\right)=\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{\pi_i^2(1-\pi_i)^2 t_i^2 e^{-2h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}\varphi_i(p_0,p_1)\right\}.
$$

Proof. $\Delta_{22}^{(i)}$ can be written as follows:

$$\Delta_{22}^{(i)} = E_O\left(\mathbf{I}_{22}^{(i)}\right) - E_O\left(\mathbf{J}_{22}^{(i)}\right)$$

$$= -E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[(1-\delta_i)\frac{v_i\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(v_i\pi_i e^{-h_it_i}+1-\pi_i)^2}\right]\right\} + E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[(1-\delta_i)\frac{\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(\pi_i e^{-h_it_i}+1-\pi_i)^2}\right]\right\}$$

$$+ E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\left[(1-\delta_i)\frac{v_i\pi_i t_i e^{-h_it_i}}{v_i\pi_i e^{-h_it_i}+1-\pi_i}\right]\right\} - E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\left[(1-\delta_i)\frac{\pi_i t_i e^{-h_it_i}}{\pi_i e^{-h_it_i}+1-\pi_i}\right]\right\}.$$

By using (20) and (21), we can write the third term in the above expression as follows:

$$E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\left[(1-\delta_i)\frac{v_i\pi_i t_i e^{-h_it_i}}{v_i\pi_i e^{-h_it_i}+1-\pi_i}\right]\right\}$$

$$= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\frac{(1-\delta_i)(1-d_i)v_i\pi_i t_i e^{-h_it_i}}{v_i\pi_i e^{-h_it_i}+1-\pi_i}\right]\right\} + E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\frac{(1-\delta_i)d_i v_i\pi_i t_i e^{-h_it_i}}{v_i\pi_i e^{-h_it_i}+1-\pi_i}\right]\right\}$$

$$= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\frac{(1-\delta_i)(1-d_i)(1-p_1)\pi_i t_i e^{-h_it_i}}{(1-p_1)\pi_i e^{-h_it_i}+(1-p_0)(1-\pi_i)}\right]\right\} + E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\frac{(1-\delta_i)d_i p_1\pi_i t_i e^{-h_it_i}}{p_1\pi_i e^{-h_it_i}+p_0(1-\pi_i)}\right]\right\}$$

$$= E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\frac{P(\delta_i=0,d_i=0|t_i)(1-p_1)\pi_i t_i e^{-h_it_i}}{(1-p_1)\pi_i e^{-h_it_i}+(1-p_0)(1-\pi_i)}\right\} + E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\frac{P(\delta_i=0,d_i=1|t_i)p_1\pi_i t_i e^{-h_it_i}}{p_1\pi_i e^{-h_it_i}+p_0(1-\pi_i)}\right\}$$

$$= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}(1-p_1)\pi_i t_i e^{-h_it_i}\right]\right\} + E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}p_1\pi_i t_i e^{-h_it_i}\right]\right\}$$

$$= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\pi_i t_i e^{-h_it_i}\right]\right\}.$$

Because the above expression does not depend on v_i, we have

$$E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\left[(1-\delta_i)\frac{v_i\pi_i t_i e^{-h_it_i}}{v_i\pi_i e^{-h_it_i}+1-\pi_i}\right]\right\} = E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{1}{h_i}\left[(1-\delta_i)\frac{\pi_i t_i e^{-h_it_i}}{\pi_i e^{-h_it_i}+1-\pi_i}\right]\right\}.$$

Therefore, it follows that

$$\Delta_{22}^{(i)} = -E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[(1-\delta_i)\frac{v_i\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(v_i\pi_i e^{-h_it_i}+1-\pi_i)^2}\right]\right\} + E_O\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\left[(1-\delta_i)\frac{\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(\pi_i e^{-h_it_i}+1-\pi_i)^2}\right]\right\}$$

$$= -E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{(1-\delta_i)(1-d_i)v_i\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(v_i\pi_i e^{-h_it_i}+1-\pi_i)^2}\right]\right\}$$

$$- E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{(1-\delta_i)d_i v_i\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(v_i\pi_i e^{-h_it_i}+1-\pi_i)^2}\right]\right\} + E_{\mathbf{T}}\left\{E_{\mathbf{V}}\left[\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{(1-\delta_i)\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(\pi_i e^{-h_it_i}+1-\pi_i)^2}\right]\right\}$$

$$= -E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{P(\delta_i=0,d_i=0|t_i)(1-p_1)(1-p_0)\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{[(1-p_1)\pi_i e^{-h_it_i}+(1-\pi_i)(1-p_0)]^2}\right\}$$

$$- E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{P(\delta_i=0,d_i=1|t_i)p_1 p_0\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{[p_1\pi_i e^{-h_it_i}+p_0(1-\pi_i)]^2}\right\} + E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\frac{P(\delta_i=0|t_i)\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(\pi_i e^{-h_it_i}+1-\pi_i)^2}\right\}.$$

Based on (20) and (21), we have

$$\Delta_{22}^{(i)} = -\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{(1-p_1)(1-p_0)\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(1-p_1)\pi_i e^{-h_it_i}+(1-\pi_i)(1-p_0)}\right\}$$

$$- \frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{p_1 p_0\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{p_1\pi_i e^{-h_it_i}+p_0(1-\pi_i)}\right\} + \frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{\pi_i e^{-h_it_i}+1-\pi_i}\right\}$$

$$= -\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{(1-p_1)(1-p_0)\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{(1-p_1)\pi_i e^{-h_it_i}+(1-\pi_i)(1-p_0)} - \frac{(1-p_0)\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{\pi_i e^{-h_it_i}+1-\pi_i}\right\}$$

$$- \frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}E_{\mathbf{T}}\left\{\frac{p_1 p_0\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{p_1\pi_i e^{-h_it_i}+p_0(1-\pi_i)} - \frac{p_0\pi_i(1-\pi_i)t_i^2 e^{-h_it_i}}{\pi_i e^{-h_it_i}+1-\pi_i}\right\}$$

$$
\begin{aligned}
&= \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'} E_{\mathbf{T}} \left\{ \frac{(1-p_0)\pi_i(1-\pi_i)t_i^2 e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \frac{(p_1 - p_0)(1-\pi_i)}{(1-p_1)\pi_i e^{-h_i t_i} + (1-\pi_i)(1-p_0)} \right\} \\
&\quad + \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'} E_{\mathbf{T}} \left\{ \frac{p_0 \pi_i(1-\pi_i)t_i^2 e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \frac{(p_0 - p_1)(1-\pi_i)}{p_1 \pi_i e^{-h_i t_i} + p_0(1-\pi_i)} \right\} \\
&= \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'} E_{\mathbf{T}} \left\{ \frac{\pi_i^2(1-\pi_i)^2 t_i^2 e^{-2h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \varphi_i(p_0, p_1) \right\}.
\end{aligned}
$$

$\square$

Because the expressions of $\Delta_{12}^{(i)}$, $\Delta_{11}^{(i)}$, and $\Delta_{22}^{(i)}$ all involve $\varphi_i(p_0, p_1)$, to prove Proposition 1, we need the following lemma regrading $\varphi_i(p_0, p_1)$.

Lemma 5 *For function*

$$
\varphi_i(p_0, p_1) = \frac{(p_0 - p_1)^2}{[(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)][p_1 \pi_i e^{-h_i t_i} + p_0(1-\pi_i)]},
$$

if $0 \le p_1 \le p_0 \le 1$, then for any i, $\varphi_i(p_0, p_1)$ is an increasing function of p_0, and a decreasing function of p_1.

Proof. If holding p_0 fixed, we can rewrite $\varphi_i(p_0, p_1)$ as

$$
\varphi_i(p_0, p_1) = \frac{p_0 - p_1}{\pi_i e^{-h_i t_i}} \left[\frac{p_0}{p_1 \pi_i e^{-h_i t_i} + p_0(1-\pi_i)} - \frac{1 - p_0}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)} \right].
$$

Because $p_0 \ge p_1$, smaller p_1 leads to larger $p_0 - p_1$, larger $\frac{p_0}{p_1 \pi_i e^{-h_i t_i} + p_0(1-\pi_i)}$, and smaller $\frac{1-p_0}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)}$. All these lead to a larger $\varphi_i(p_0, p_1)$. If we hold p_1 as fixed, $\varphi_i(p_0, p_1)$ can be rewritten as

$$
\varphi_i(p_0, p_1) = \frac{p_0 - p_1}{1 - \pi_i} \left[\frac{1 - p_1}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)} - \frac{p_1}{p_1 \pi_i e^{-h_i t_i} + p_0(1-\pi_i)} \right].
$$

Larger p_0 leads to larger $p_0 - p_1$, larger $\frac{1-p_1}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)}$, and smaller $\frac{p_1}{p_1 \pi_i e^{-h_i t_i} + p_0(1-\pi_i)}$. These lead to a larger $\varphi_i(p_0, p_1)$. $\square$

With the differences for each entry of the information matrix computed by Lemmas 2 - 4, and the property of the differences established by Lemma 5, we are ready to prove Proposition 1.

Proof of Proposition 1. Let

$$
a_{11}^{(i)} = \left[\frac{\delta_i}{\pi_i^2} + (1-\delta_i) \frac{(e^{-h_i t_i} - 1)^2}{(\pi_i e^{-h_i t_i} + 1 - \pi_i)^2} \right] - \frac{1 - e^{2\gamma' \mathbf{z}_i}}{e^{\gamma' \mathbf{z}_i}} \left[\frac{\delta_i}{\pi_i} + (1-\delta_i) \frac{e^{-h_i t_i} - 1}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right],
$$

$$
a_{22}^{(i)} = \left[\frac{\delta_i}{h_i^2} - (1-\delta_i) \frac{\pi_i(1-\pi_i)t_i^2 e^{-h_i t_i}}{(\pi_i e^{-h_i t_i} + 1 - \pi_i)^2} \right] - \frac{1}{h_i} \left[\frac{\delta_i}{h_i} - \delta_i t_i - (1-\delta_i) \frac{\pi_i t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right],
$$

$$
a_{12}^{(i)} = (1-\delta_i) \frac{t_i e^{-h_i t_i}}{(\pi_i e^{-h_i t_i} + 1 - \pi_i)^2},
$$

$$
d_{11}^{(i)} = \frac{e^{-2h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i},
$$

$$
d_{22}^{(i)} = \frac{\pi_i^2(1-\pi_i)^2 t_i^2 e^{-2h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i},
$$

$$
d_{12}^{(i)} = \frac{t_i \pi_i e^{-2h_i t_i}(1-\pi_i)}{\pi_i e^{-h_i t_i} + 1 - \pi_i}.
$$

Then we have

$$
E_O(\mathbf{J}_{11}) = E_O \left(\sum_{i=1}^{n} a_{11}^{(i)} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'} \right),
$$

$$E_O(\mathbf{J}_{22}) = E_O\left(\sum_{i=1}^n a_{22}^{(i)} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right),$$

$$E_O(\mathbf{J}_{12}) = E_O\left(\sum_{i=1}^n a_{12}^{(i)} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right)$$

and, by Lemmas 2 - 4,

$$\Delta_{11} = \sum_{i=1}^n \Delta_{11}^{(i)} = E_{\mathbf{T}}\left(\sum_{i=1}^n d_{11}^{(i)} \varphi_i(p_0, p_1) \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\right),$$

$$\Delta_{22} = \sum_{i=1}^n \Delta_{22}^{(i)} = E_{\mathbf{T}}\left(\sum_{i=1}^n d_{22}^{(i)} \varphi_i(p_0, p_1) \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right),$$

$$\Delta_{12} = \sum_{i=1}^n \Delta_{12}^{(i)} = -E_{\mathbf{T}}\left(\sum_{i=1}^n d_{12}^{(i)} \varphi_i(p_0, p_1) \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right).$$

It is obvious from the above expressions that $a_{12}^{(i)} \geq 0$, $d_{12}^{(i)} \geq 0$, and $E_O[a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1)] \geq 0$. Also, because of the positive definite property of the information matrix, we have $a_{11}^{(i)} \geq 0$, $a_{22}^{(i)} \geq 0$, $d_{11}^{(i)} \geq 0$, and $d_{22}^{(i)} \geq 0$.

Proof of Case 1. When $p_0 = 1$ and $p_1 = 0$, $\varphi_i(p_0, p_1)$ reduces to

$$\varphi_i(p_0, p_1) = \frac{1}{(1 - \pi_i)\pi_i e^{-h_i t_i}}.$$

For any i, we have

$$d_{12}^{(i)} \varphi_i(p_0, p_1) = \frac{t_i \pi_i e^{-2h_i t_i}(1 - \pi_i)}{\pi_i e^{-h_i t_i} + (1 - \pi_i)} \frac{1}{(1 - \pi_i)\pi_i e^{-h_i t_i}} = \frac{t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i},$$

and by (23)

$$E_O\left(a_{12}^{(i)}\right) = E_{\mathbf{T}}\left[\frac{t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i}\right].$$

Thus

$$E_O(\mathbf{I}_{12}) = E_O\left(\sum_{i=1}^n \left[a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1)\right] \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right) = \mathbf{0}. \tag{24}$$

Using (24) we have

$$V_{\boldsymbol{\gamma}}^D = \left\{E_O(\mathbf{I}_{11}) - E_O(\mathbf{I}_{12})E_O^{-1}(\mathbf{I}_{22})E_O'(\mathbf{I}_{12})\right\}^{-1} = \left\{E_O\left(\sum_{i=1}^n \left[a_{11}^{(i)} + d_{11}^{(i)} \varphi_i(p_0, p_1)\right] \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\right)\right\}^{-1}$$

and

$$V_{\boldsymbol{\beta}}^D = \left\{E_O(\mathbf{I}_{22}) - E_O'(\mathbf{I}_{12})E_O^{-1}(\mathbf{I}_{11})E_O(\mathbf{I}_{12})\right\}^{-1} = \left\{E_O\left(\sum_{i=1}^n \left[a_{22}^{(i)} + d_{22}^{(i)} \varphi_i(p_0, p_1)\right] \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right)\right\}^{-1}.$$

For any $(k + 1)$ dimensional vectors $\boldsymbol{u}$ and $\boldsymbol{w}$, a non-negative constant c, and a $(k + 1) \times (k + 1)$ positive definite matrix $\boldsymbol{A}$, we have

$$\boldsymbol{u}'(\boldsymbol{A} + c\boldsymbol{w}\boldsymbol{w}')^{-1}\boldsymbol{u} = \boldsymbol{u}'\left(\boldsymbol{A}^{-1} - \frac{1}{\frac{1}{c} + \boldsymbol{w}'\boldsymbol{A}^{-1}\boldsymbol{w}} \boldsymbol{A}^{-1}\boldsymbol{w}\boldsymbol{w}'\boldsymbol{A}^{-1}\right)\boldsymbol{u}$$

$$= \boldsymbol{u}'\boldsymbol{A}^{-1}\boldsymbol{u} - \frac{1}{\frac{1}{c} + \boldsymbol{w}'\boldsymbol{A}^{-1}\boldsymbol{w}} \boldsymbol{u}^T \boldsymbol{A}^{-1}\boldsymbol{w}\boldsymbol{w}'\boldsymbol{A}^{-1}\boldsymbol{u}$$

$$= \boldsymbol{u}^T \boldsymbol{A}^{-1}\boldsymbol{u} - \frac{1}{\frac{1}{c} + \boldsymbol{w}'\boldsymbol{A}^{-1}\boldsymbol{w}} (\boldsymbol{u}'\boldsymbol{A}^{-1}\boldsymbol{w})^2 \leq \boldsymbol{u}'\boldsymbol{A}^{-1}\boldsymbol{u}. \tag{25}$$

By adding $\Delta_{11}^{(i)}$ one at a time, for any $\boldsymbol{u}$, we have $\boldsymbol{u}' V_{\boldsymbol{\gamma}}^D \boldsymbol{u} \leq \boldsymbol{u}' V_{\boldsymbol{\gamma}}^N \boldsymbol{u}$. By taking $\boldsymbol{u}_i$ as $u_{ij} = 0$ if $j \neq i$ and $u_{ij} = 1$ if $j = i$, we can conclude that all the diagonal entries of $V_{\boldsymbol{\gamma}}^D$ are less than or equal to the corresponding diagonal entries of $V_{\boldsymbol{\gamma}}^N$. Because smaller diagonal entries indicate higher efficiency, the estimator of $\boldsymbol{\gamma}$ with diagnostic information included is more efficient than that without diagnostic information included.

Similarly, it can be shown that all the diagonal entries of $V_{\boldsymbol{\beta}}^D$ are less than or equal to the corresponding diagonal entries of $V_{\boldsymbol{\beta}}^N$ and, hence, the estimate of $\boldsymbol{\beta}$ with diagnostic information included is more efficient than that without diagnostic information included. $\qquad \square$

Proof of Case 2. Because $\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}$ and $\frac{\partial h_i}{\partial \boldsymbol{\beta}}$ are the same for all subjects when $\boldsymbol{\gamma} = (\gamma_0)$ and $\boldsymbol{\beta} = (\beta_0)$, denote

$$\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} = C_{\gamma_0} \text{ and } \frac{\partial h_i}{\partial \boldsymbol{\beta}} = C_{\beta_0}.$$

Then we have

$$E_{\boldsymbol{O}}(\mathbf{I}_{11}) = E_{\boldsymbol{O}} \left(\sum_{i=1}^n \left[a_{11}^{(i)} + d_{11}^{(i)} \varphi_i(p_0, p_1) \right] \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'} \right) = n C_{\gamma_0}^2 E_{\boldsymbol{O}} \left(a_{11}^{(1)} + d_{11}^{(1)} \varphi_1(p_0, p_1) \right), \tag{26}$$

$$E_{\boldsymbol{O}}(\mathbf{I}_{12}) = E_{\boldsymbol{O}} \left(\sum_{i=1}^n \left[a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1) \right] \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'} \right) = n C_{\gamma_0} C_{\beta_0} E_{\boldsymbol{O}} \left(a_{12}^{(1)} - d_{12}^{(1)} \varphi_1(p_0, p_1) \right), \tag{27}$$

and

$$E_{\boldsymbol{O}}(\mathbf{I}_{22}) = E_{\boldsymbol{O}} \left(\sum_{i=1}^n \left[a_{22}^{(i)} + d_{22}^{(i)} \varphi_i(p_0, p_1) \right] \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}'} \right) = n C_{\beta_0}^2 E_{\boldsymbol{O}} \left(a_{22}^{(1)} + d_{22}^{(1)} \varphi_1(p_0, p_1) \right). \tag{28}$$

From (26), (27), and (28), we have

$$\left[V_{\boldsymbol{\gamma}}^D \right]^{-1} = E_{\boldsymbol{O}}(\mathbf{I}_{11}) - E_{\boldsymbol{O}}(\mathbf{I}_{12}) E_{\boldsymbol{O}}^{-1}(\mathbf{I}_{22}) E_{\boldsymbol{O}}'(\mathbf{I}_{12}) = n C_{\gamma_0}^2 E_{\boldsymbol{O}} \left(a_{11}^{(1)} + d_{11}^{(1)} \varphi_1(p_0, p_1) \right) - n C_{\gamma_0}^2 \frac{E_{\boldsymbol{O}}^2 (a_{12}^{(1)} - d_{12}^{(1)} \varphi_1(p_0, p_1))}{E_{\boldsymbol{O}}(a_{22}^{(1)} + d_{22}^{(1)} \varphi_1(p_0, p_1))}.$$

Because $E_{\boldsymbol{O}}(a_{11}^{(1)} + d_{11}^{(1)} \varphi_1(p_0, p_1))$ and $E_{\boldsymbol{O}}(a_{22}^{(1)} + d_{22}^{(1)} \varphi_1(p_0, p_1))$ are increasing functions of p_0, and decreasing functions of p_1 through their dependence on $\varphi_1(p_0, p_1)$, and $E_{\boldsymbol{O}}^2(a_{12}^{(1)} - d_{12}^{(1)} \varphi_1(p_0, p_1))$ is a decreasing function of p_0, and an increasing function of p_1 through its dependence on $\varphi_1(p_0, p_1)$, $\left[V_{\boldsymbol{\gamma}}^D \right]^{-1}$ is an increasing function of p_0, and a decreasing function of p_1, i.e., an increasing function of p_0 (sensitivity) and $1 - p_1$ (specificity). Larger $\left[V_{\boldsymbol{\gamma}}^D \right]^{-1}$ leads to smaller $V_{\boldsymbol{\gamma}}^D$. Thus the efficiency of the estimator of $\boldsymbol{\gamma}$ increases as either specificity or sensitivity increases, and the estimator of $\boldsymbol{\gamma}$ with diagnostic information included is more efficient than that without diagnostic information included.

Similarly, we have

$$\left[V_{\boldsymbol{\beta}}^D \right]^{-1} = E_{\boldsymbol{O}}(\mathbf{I}_{22}) - E_{\boldsymbol{O}}'(\mathbf{I}_{12}) E_{\boldsymbol{O}}^{-1}(\mathbf{I}_{11}) E_{\boldsymbol{O}}(\mathbf{I}_{12})$$
$$= n C_{\beta_0}^2 E_{\boldsymbol{O}} \left(a_{22}^{(1)} + d_{22}^{(1)} \varphi_1(p_0, p_1) \right) - n C_{\beta_0}^2 \frac{E_{\boldsymbol{O}}^2 (a_{12}^{(1)} - d_{12}^{(1)} \varphi_1(p_0, p_1))}{E_{\boldsymbol{O}}(a_{11}^{(1)} + d_{11}^{(1)} \varphi_1(p_0, p_1))}.$$

$\left[V_{\boldsymbol{\beta}}^D \right]^{-1}$ is also an increasing function of p_0, and a decreasing function of p_1, i.e., an increasing function of p_0 (sensitivity) and $1 - p_1$ (specificity). Larger $\left[V_{\boldsymbol{\beta}}^D \right]^{-1}$ leads to smaller $V_{\boldsymbol{\beta}}^D$. Therefore, the efficiency of the estimator of $\boldsymbol{\beta}$ increases as either specificity or sensitivity increases, and the estimator of $\boldsymbol{\beta}$ with diagnostic information included is more efficient than that without diagnostic information included. $\qquad \square$

Proof of Case 3. $\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}$ is the same for all subjects when $\boldsymbol{\gamma} = (\gamma_0)$, so we can denote it as an unknown constant C_{γ_0}. For $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ and x_{i1} being a binary variable with values of 0 and 1, $\frac{\partial h_i}{\partial \boldsymbol{\beta}}$ can be expressed as follows:

$$\frac{\partial h_i}{\partial \boldsymbol{\beta}} = h_i \begin{bmatrix} 1 \\ x_{i1} \end{bmatrix} = e_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} I(x_{i1} = 0) + e_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} I(x_{i1} = 1).$$

Here $e_0 = \exp(\beta_0)$ and $e_1 = \exp(\beta_0 + \beta_1)$ are the shorthand notations of $h_i = h(x_{i1} = 0)$ and $h_i = h(x_{i1} = 1)$, respectively. $I(\cdot)$ is an indicator function.

For $j = 0, 1$, let

$$b_{11j} = E_O\left(a_{11}^{(i)} + d_{11}^{(i)}\varphi_i(p_0, p_1)|x_{i1} = j\right),$$
$$b_{12j} = E_O\left(a_{12}^{(i)} - d_{12}^{(i)}\varphi_i(p_0, p_1)|x_{i1} = j\right),$$
$$b_{22j} = E_O\left(a_{22}^{(i)} + d_{22}^{(i)}\varphi_i(p_0, p_1)|x_{i1} = j\right).$$

Assume that there are n_0 observations with $x_{i1} = 0$, and n_1 observations with $x_{i1} = 1$. By the independent and identically distributed (i.i.d.) property when the covariates are the same, we have

$$E_O(\mathbf{I}_{11}) = E_O\left(\sum_{i=1}^{n}\left[a_{11}^{(i)} + d_{11}^{(i)}\varphi_i(p_0, p_1)\right]\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}'}\right) = C_{\gamma_0}^2(n_1 b_{111} + n_0 b_{110}), \tag{29}$$

$$\begin{aligned}
E_O(\mathbf{I}_{12}) &= E_O\left(\sum_{i=1}^{n}\left[a_{12}^{(i)} - d_{12}^{(i)}\varphi_i(p_0, p_1)\right]\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right) \\
&= C_{\gamma_0}[n_1 b_{121}e_1(1, 1) + n_0 b_{120}e_0(1, 0)] \\
&= C_{\gamma_0}(n_1 b_{121}e_1 + n_0 b_{120}e_0, n_1 b_{121}e_1),
\end{aligned} \tag{30}$$

and

$$\begin{aligned}
E_O(\mathbf{I}_{22}) &= E_O\left(\sum_{i=1}^{n}\left[a_{22}^{(i)} + d_{22}^{(i)}\varphi_i(p_0, p_1)\right]\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}'}\right) \\
&= n_1 b_{221}e_1^2\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + n_0 b_{220}e_0^2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} n_1 b_{221}e_1^2 + n_0 b_{220}e_0^2 & n_1 b_{221}e_1^2 \\ n_1 b_{221}e_1^2 & n_1 b_{221}e_1^2 \end{bmatrix}.
\end{aligned}$$

As a result, the inverse of $E_O(\mathbf{I}_{22})$ is

$$E_O^{-1}(\mathbf{I}_{22}) = \begin{bmatrix} \frac{1}{n_0 b_{220}e_0^2} & -\frac{1}{n_0 b_{220}e_0^2} \\ -\frac{1}{n_0 b_{220}e_0^2} & \frac{1}{n_0 b_{220}e_0^2} + \frac{1}{n_1 b_{221}e_1^2} \end{bmatrix}. \tag{31}$$

From (29), (30), and (31), the inverse of V_γ^D is as follows:

$$\begin{aligned}
\left[V_\gamma^D\right]^{-1} &= E_O(\mathbf{I}_{11}) - E_O(\mathbf{I}_{12})E_O^{-1}(\mathbf{I}_{22})E_O'(\mathbf{I}_{12}) \\
&= C_{\gamma_0}^2(n_1 b_{111} + n_0 b_{110}) \\
&\quad - C_{\gamma_0}^2(n_1 b_{121}e_1 + n_0 b_{120}e_0, n_1 b_{121}e_1)\begin{bmatrix} \frac{1}{n_0 b_{220}e_0^2} & -\frac{1}{n_0 b_{220}e_0^2} \\ -\frac{1}{n_0 b_{220}e_0^2} & \frac{1}{n_0 b_{220}e_0^2} + \frac{1}{n_1 b_{221}e_1^2} \end{bmatrix}\begin{bmatrix} n_1 b_{121}e_1 + n_0 b_{120}e_0 \\ n_1 b_{121}e_1 \end{bmatrix} \\
&= C_{\gamma_0}^2\left[n_1 b_{111} + n_0 b_{110} - \frac{(n_0 b_{120}e_0)^2}{n_0 b_{220}e_0^2} - \frac{(n_1 b_{121}e_1)^2}{n_1 b_{221}e_1^2}\right] \\
&= C_{\gamma_0}^2\left(n_1 b_{111} + n_0 b_{110} - \frac{n_0 b_{120}^2}{b_{220}} - \frac{n_1 b_{121}^2}{b_{221}}\right).
\end{aligned}$$

Because b_{111}, b_{110}, b_{220}, and b_{221} are increasing functions of p_0, and decreasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, and b_{121} and b_{120} are decreasing functions of p_0, and increasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, $\left[V_\gamma^D\right]^{-1}$ is an increasing function of p_0, and a decreasing function of p_1, i.e., an increasing function of p_0 (sensitivity) and $1 - p_1$ (specificity). Larger $\left[V_\gamma^D\right]^{-1}$ leads to smaller V_γ^D. Consequently, the efficiency of the estimator of $\boldsymbol{\gamma}$ increases as either specificity or sensitivity increases, and the estimator of $\boldsymbol{\gamma}$ with diagnostic information included is more efficient than that without diagnostic information included.

For $V_{\boldsymbol{\beta}}^{D}$, we have

$$\frac{1}{G}\left[\begin{array}{cc} V_{11} & V_{12} \\ V_{12} & V_{22} \end{array}\right],$$

where

$$G = (n_1 b_{221} e_1^2)(n_0 b_{220} e_0^2)(n_1 b_{111} + n_0 b_{110}) - (n_1 b_{221} e_1^2)(n_0 b_{120} e_0)^2 - (n_0 b_{220} e_0^2)(n_1 b_{121} e_1)^2,$$
$$V_{11} = (n_1 b_{221} e_1^2)(n_1 b_{111} + n_0 b_{110}) - (n_1 b_{121} e_1)^2,$$
$$V_{22} = (n_1 b_{221} e_1^2 + n_0 b_{220} e_0^2)(n_1 b_{111} + n_0 b_{110}) - (n_1 b_{121} e_1 + n_0 b_{120} e_0)^2,$$
$$V_{12} = -(n_1 b_{221} e_1^2)(n_1 b_{111} + n_0 b_{110}) + (n_1 b_{121} e_1 + n_0 b_{120} e_0)(n_1 b_{121} e_1).$$

For the asymptotic variance of the MLE of β_0 with diagnostic information included, we have

$$V_{\beta_0}^{D} = \frac{(n_1 b_{221} e_1^2)(n_1 b_{111} + n_0 b_{110}) - (n_1 b_{121} e_1)^2}{(n_1 b_{221} e_1^2)(n_0 b_{220} e_0^2)(n_1 b_{111} + n_0 b_{110}) - (n_1 b_{221} e_1^2)(n_0 b_{120} e_0)^2 - (n_0 b_{220} e_0^2)(n_1 b_{121} e_1)^2}$$

$$= \frac{1}{n_0 b_{220} e_0^2 - \frac{(n_1 b_{221} e_1^2)(n_0 b_{120} e_0)^2}{(n_1 b_{221} e_1^2)(n_1 b_{111} + n_0 b_{110}) - (n_1 b_{121} e_1)^2}}$$

$$= \frac{1}{n_0 b_{220} e_0^2 - \frac{(n_0 b_{120} e_0)^2}{n_1 b_{111} + n_0 b_{110} - \frac{(n_1 b_{121} e_1)^2}{n_1 b_{221} e_1^2}}}.$$

Because b_{111}, b_{110}, b_{220}, and b_{221} are increasing functions of p_0, and decreasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, and b_{121} and b_{120} are decreasing functions of p_0, and increasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, $\frac{(n_1 b_{121} e_1)^2}{n_1 b_{221} e_1^2}$ is a decreasing function of p_0 and an increasing function of p_1 and, hence, $n_1 b_{111} + n_0 b_{110} - \frac{(n_1 b_{121} e_1)^2}{n_1 b_{221} e_1^2}$ is an increasing function of p_0 and a decreasing function of p_1. This implies that $\frac{(n_0 b_{120} e_0)^2}{n_1 b_{111} + n_0 b_{110} - \frac{(n_1 b_{121} e_1)^2}{n_1 b_{221} e_1^2}}$ is a decreasing function of p_0 and an increasing function of p_1, $n_0 b_{220} e_0^2 - \frac{(n_0 b_{120} e_0)^2}{n_1 b_{111} + n_0 b_{110} - \frac{(n_1 b_{121} e_1)^2}{n_1 b_{221} e_1^2}}$ is an increasing function of p_0 and a decreasing function of p_1, and $V_{\beta_0}^{D}$ is a decreasing function of p_0 and an increasing function of p_1, i.e., a decreasing function of p_0 (sensitivity) and $1 - p_1$ (specificity). Therefore, the efficiency of the estimator of β_0 increases as either specificity or sensitivity increases. Because the estimator of β_0 without diagnostic information included corresponds to the case here sensitivity is the same as 1 - specificity ($p_0 = p_1$), the estimator of β_0 with diagnostic information included (with $p_0 > p_1$) is more efficient than that without diagnostic information included. □

Proof of Case 4. When $\boldsymbol{\beta} = (\beta_0)$, because $\frac{\partial h_i}{\partial \beta}$ is the same for all subjects, it is denoted as an unknown constant C_{β_0}. For $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)'$, we can express $\frac{\partial \pi_i}{\partial \gamma}$ as

$$\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} = \frac{e^{\boldsymbol{\gamma}' \mathbf{z}_i}}{(1 + e^{\boldsymbol{\gamma}' \mathbf{z}_i})^2} \left[\begin{array}{c} 1 \\ z_{i1} \end{array}\right] = g_0 \left[\begin{array}{c} 1 \\ 0 \end{array}\right] I(z_{i1} = 0) + g_1 \left[\begin{array}{c} 1 \\ 1 \end{array}\right] I(z_{i1} = 1),$$

where $g_0 = \frac{e^{\gamma_0}}{(1 + e^{\gamma_0})^2}$ when $z_{i1} = 0$ and $g_1 = \frac{e^{\gamma_0 + \gamma_1}}{(1 + e^{\gamma_0 + \gamma_1})^2}$ when $z_{i1} = 1$.
For $j = 0, 1$, let

$$c_{11j} = E_{\boldsymbol{O}}\left(a_{11}^{(i)} + d_{11}^{(i)} \varphi_i(p_0, p_1) | z_{i1} = j\right),$$
$$c_{12j} = E_{\boldsymbol{O}}\left(a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1) | z_{i1} = j\right),$$
$$c_{22j} = E_{\boldsymbol{O}}\left(a_{22}^{(i)} + d_{22}^{(i)} \varphi_i(p_0, p_1) | z_{i1} = j\right).$$

Suppose there are n_0^* observations with $z_{i1} = 0$ and n_1^* observations with $z_{i1} = 1$. It then can be obtained with the

i.i.d. property when the covariates are the same that

$$
E_O(\mathbf{I}_{11}) = E_O\left(\sum_{i=1}^{n}\left[a_{11}^{(i)} + d_{11}^{(i)}\varphi_i(p_0, p_1)\right]\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}'}\right)
$$

$$
= n_1^* c_{111} g_1^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + n_0^* c_{110} g_0^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}
$$

$$
= \begin{bmatrix} n_1^* c_{111} g_1^2 + n_0^* c_{110} g_0^2 & n_1^* c_{111} g_1^2 \\ n_1^* c_{111} g_1^2 & n_1^* c_{111} g_1^2 \end{bmatrix},
$$

which implies that

$$
E_O^{-1}(\mathbf{I}_{11}) = \begin{bmatrix} \frac{1}{n_0^* c_{110} g_0^2} & -\frac{1}{n_0^* c_{110} g_0^2} \\ -\frac{1}{n_0^* c_{110} g_0^2} & \frac{1}{n_0^* c_{110} g_0^2} + \frac{1}{n_1^* c_{111} g_1^2} \end{bmatrix}. \tag{32}
$$

Because

$$
E_O(\mathbf{I}_{12}) = E_O\left\{\sum_{i=1}^{n}\left[a_{12}^{(i)} - d_{12}^{(i)}\varphi_i(p_0, p_1)\right]\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial h_i}{\partial\boldsymbol{\beta}'}\right\}
$$

$$
= C_{\beta_0}\left\{n_1^* c_{121} g_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + n_0^* c_{120} g_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}
$$

$$
= C_{\beta_0}\begin{bmatrix} n_1^* c_{121} g_1 + n_0^* c_{120} g_0 \\ n_1^* c_{121} g_1 \end{bmatrix} \tag{33}
$$

and

$$
E_O(\mathbf{I}_{22}) = E_O\left(\sum_{i=1}^{n}\left[a_{22}^{(i)} + d_{22}^{(i)}\varphi_i(p_0, p_1)\right]\frac{\partial h_i}{\partial\boldsymbol{\beta}}\frac{\partial h_i}{\partial\boldsymbol{\beta}'}\right) = C_{\beta_0}^2(n_1^* c_{221} + n_0^* c_{220}). \tag{34}
$$

From (32), (33), and (34), we have

$$
\left[V_\beta^D\right]^{-1} = E_O(\mathbf{I}_{22}) - E_O'(\mathbf{I}_{12})E_O^{-1}(\mathbf{I}_{11})E_O(\mathbf{I}_{12})
$$

$$
= C_{\beta_0}^2(n_1^* c_{221} + n_0^* c_{220})
$$

$$
- C_{\beta_0}^2(n_1^* c_{121} g_1 + n_0^* c_{120} g_0, n_1^* c_{121} g_1)\begin{bmatrix} \frac{1}{n_0^* c_{110} g_0^2} & -\frac{1}{n_0^* c_{110} g_0^2} \\ -\frac{1}{n_0^* c_{110} g_0^2} & \frac{1}{n_0^* c_{110} g_0^2} + \frac{1}{n_1^* c_{111} g_1^2} \end{bmatrix}\begin{bmatrix} n_1^* c_{121} g_1 + n_0^* c_{120} g_0 \\ n_1^* c_{121} g_1 \end{bmatrix}
$$

$$
= C_{\beta_0}^2\left[n_1^* c_{221} + n_0^* c_{220} - \frac{(n_0^* c_{120} g_0)^2}{n_0^* c_{110} g_0^2} - \frac{(n_1^* c_{121} g_1)^2}{n_1^* c_{111} g_1^2}\right]
$$

$$
= C_{\beta_0}^2\left(n_1^* c_{221} + n_0^* c_{220} - \frac{n_0^* c_{120}^2}{c_{110}} - \frac{n_1^* c_{121}^2}{c_{111}}\right).
$$

c_{111}, c_{110}, c_{220}, and c_{221} are increasing functions of p_0, and decreasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, and c_{121} and c_{120} are decreasing functions of p_0, and increasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, so $\left[V_\beta^D\right]^{-1}$ is an increasing function of p_0, and a decreasing function of p_1, i.e., an increasing function of p_0 (sensitivity) and $1 - p_1$ (specificity). Larger $\left[V_\beta^D\right]^{-1}$ leads to smaller V_β^D. Consequently, the efficiency of the estimator of $\boldsymbol{\beta}$ increases as either specificity or sensitivity increases, and the estimator of $\boldsymbol{\beta}$ with diagnostic information included is more efficient than that without diagnostic information included.

For V_γ^D, it can be computed as

$$
\frac{1}{G^*}\begin{bmatrix} V_{11}^* & V_{12}^* \\ V_{12}^* & V_{22}^* \end{bmatrix},
$$

where

$$
G^* = (n_1^* c_{111} g_1^2)(n_0^* c_{110} g_0^2)(n_1^* c_{221} + n_0^* c_{220}) - (n_1^* c_{111} g_1^2)(n_0^* c_{120} g_0)^2 - (n_0^* c_{110} g_0^2)(n_1^* c_{121} g_1)^2,
$$

$$
V_{11}^* = (n_1^* c_{111} g_1^2)(n_1^* c_{221} + n_0^* c_{220}) - (n_1^* c_{121} g_1)^2,
$$

$$
V_{22}^* = (n_1^* c_{111} g_1^2 + n_0^* c_{110} g_0^2)(n_1^* c_{221} + n_0^* c_{220}) - (n_1^* c_{121} g_1 + n_0^* c_{120} g_0)^2,
$$

$$
V_{12}^* = -(n_1^* c_{111} g_1^2)(n_1^* c_{221} + n_0^* c_{220}) + (n_1^* c_{121} g_1 + n_0^* c_{120} g_0)(n_1^* c_{121} g_1),
$$

so the asymptotic variance of the MLE of γ_0 with diagnostic information included is as follows:

$$V_{\gamma_0}^D = \frac{(n_1^* c_{111} g_1^2)(n_1^* c_{221} + n_0^* c_{220}) - (n_1^* c_{121} g_1)^2}{(n_1^* c_{111} g_1^2)(n_0^* c_{110} g_0^2)(n_1^* c_{221} + n_0^* c_{220}) - (n_1^* c_{111} g_1^2)(n_0^* c_{120} g_0)^2 - (n_0^* c_{110} g_0^2)(n_1^* c_{121} g_1)^2}$$

$$= \frac{1}{n_0^* c_{110} g_0^2 - \frac{(n_1^* c_{111} g_1^2)(n_0^* c_{120} g_0)^2}{(n_1^* c_{111} g_1^2)(n_1^* c_{221} + n_0^* c_{220}) - (n_1^* c_{121} g_1)^2}}$$

$$= \frac{1}{n_0^* c_{110} g_0^2 - \frac{(n_0^* c_{120} g_0)^2}{n_1^* c_{221} + n_0^* c_{220} - \frac{(n_1^* c_{121} g_1)^2}{n_1^* c_{111} g_1^2}}}.$$

Because c_{111}, c_{110}, c_{220}, and c_{221} are increasing functions of p_0, and decreasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, and c_{121} and c_{120} are decreasing functions of p_0, and increasing functions of p_1 through their dependence on $\varphi_i(p_0, p_1)$, we can draw an inference as follows: $\frac{(n_1^* c_{121} g_1)^2}{n_1^* c_{111} g_1^2}$ is a decreasing function of p_0 and an increasing function of p_1 and, hence, $n_1^* c_{221} + n_0^* c_{220} - \frac{(n_1^* c_{121} g_1)^2}{n_1^* c_{111} g_1^2}$ is an increasing function of p_0 and a decreasing function of p_1. This can imply that $\frac{(n_0^* c_{120} g_0)^2}{n_1^* c_{221} + n_0^* c_{220} - \frac{(n_1^* c_{121} g_1)^2}{n_1^* c_{111} g_1^2}}$ is a decreasing function of p_0 and an increasing function of p_1, and $n_0^* c_{110} g_0^2 - \frac{(n_0^* c_{120} g_0)^2}{n_1^* c_{221} + n_0^* c_{220} - \frac{(n_1^* c_{121} g_1)^2}{n_1^* c_{111} g_1^2}}$ is an increasing function of p_0 and a decreasing function of p_1. It turns out that $V_{\gamma_0}^D$ is a decreasing function of p_0 and an increasing function of p_1, i.e., a decreasing function of p_0 (sensitivity) and $1 - p_1$ (specificity). Therefore, the efficiency of the estimator of γ_0 increases as either specificity or sensitivity increases. Because the estimate of γ_0 without diagnostic information included corresponds to the case where sensitivity is the same as 1 - specificity ($p_0 = p_1$), the estimator of γ_0 with diagnostic information included (with $p_0 > p_1$) is more efficient than that without diagnostic information included. $\square$

4. Summary and Discussion

An extended cure model incorporated with additional diagnostic information about cured status is very useful to model the failure time data where some individuals could eventually experience, and others never experience, the event of interest when their diagnostic information is available. In this paper, we have shown theoretically that the MLEs for the parameters in the extended exponential cure model are asymptotically more efficient than the MLEs for those in the classical exponential cure model. Specifically we showed for some special cases that the asymptotic efficiency increases as the sensitivity and the specificity of diagnostic procedures increase. In conclusion, based on the results provided in this paper, we highly recommend that when additional cure information is available, even only partially, we should incorporate this information into the model. It is also recommended that investigators should devise diagnostic procedures of cure and collect available cure information when we design and conduct studies.

References

Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, 11*, 15-53.

Cantor, A. B., & Shuster, J. J. (1992). Parametric versus non-parametric methods for estimating cure rates based on censored survival data. *Statistics in Medicine, 11*, 931-937. http://dx.doi.org/10.1002/sim.4780110710

Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics, 38*, 1041-1046. http://dx.doi.org/10.2307/2529885

Farewell, V. T. (1986). Mixture models in survival analysis: are they worth the risk? *Canad. J. Statist., 14*, 257-262.

Gamel, J. W., McLean, I. W., & Rosenberg, S. H. (1990). Proportion cured and mean log survival time as functions of tumor size. *Statistics in Medicine, 9*, 999-1006. http://dx.doi.org/10.1002/sim.4780090814

Ghitany, M. E., & Maller, R. A. (1992). Asymptotic results for exponential mixture models with long term sur-vivors. *Statistics, 23*, 321-336. http://dx.doi.org/10.1080/02331889208802379

Goldman, A. I. (1984). Survivorship analysis when cure is a possibility: A monte carlo study. *Statistics in Medicine, 3*, 153-163. http://dx.doi.org/10.1002/sim.4780030208

Gordon, N. H. (1990a). Maximum likelihood estimation for mixtures of two gompertz distributions when censor-ing occurs. *Communications in Statistics-Simulation and Computation, 19*, 733-747. http://dx.doi.org/10.1080/03610919008812885

Gordon, N. H. (1990b). Application of the theory of finite mixtures for the estimation of cure rates of treated cancer patients. *Statistics in Medicine, 9*, 397-407. http://dx.doi.org/10.1002/sim.4780090411

Jones, D. R., Powles, R. L., Machin, D., & Sylvester, R. J. (1981). On estimating the proportion of cured patients in clinical studies. *Biometrie-Praximetrie, 21*, 1-11.

Kuk, A. Y. C., & Chen, C. (1992). A mixture model combining logistic regression with proportional hazards regression. *Biometrika, 79*, 531-541. http://dx.doi.org/10.2307/2336784

Li, C. S., & Taylor, J. M. G. (2002). A semi-parametric accelerated failure time cure model. *Statistics in Medicine, 21*, 3235-3247. http://dx.doi.org/10.1002/sim.1260

Peng, Y., Dear, K. B. G., & Denham, J. W. (1998). A generalized F mixture model for cure rate estimation. *Statistics in Medicine, 17*, 813-830. http://dx.doi.org/10.1002/(SICI)1097-0258(19980430)17:8<813::AID-SIM775>3.0.CO;2-#

Peng, Y., & Dear, K. B. G. (2000). A nonparametric mixture model for cure rate estimation. *Biometrics, 56*, 237-243.

Sy, J. P., & Taylor, J. M. G. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics, 56*, 227-236.

Taylor, J. M. G. (1995). Semi-parametric estimation in failure time mixture models. *Biometrics, 51*, 899-907. http://dx.doi.org/10.2307/2532991

Wu, Y., Lin, Y., Lu, S. E., Li, C. S., & Shih, W. J. (2014). Extension of a Cox proportional hazards cure model when cure information is partially known. To appear in *Biostatistics*. http://dx.doi.org/10.1093/biostatistics/kxu002

Yamaguchi, K. (1992). Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of "permanent employment" in Japan. *Journal of the American Statistical Association, 87*, 284-292. http://dx.doi.org/10.2307/2290258

Zhang, J., & Peng, Y. (2007). A new estimation method for the semiparametric accelerated failure time mixture cure model. *Statistics in Medicine, 26*(16), 3157-3171. http://dx.doi.org/10.1002/sim.2748

Copyrights