EXTENSION OF CURE RATE MODEL WHEN CURED IS PARTIALLY KNOWN

BY YU WU

A dissertation submitted to the

School of Public Health

University of Medicine and Dentistry of New Jersey

and the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

UMDNJ-School of Public Health

Awarded jointly by these institutions and

Written under the direction of

Yong Lin, Ph.D

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2010

ABSTRACT OF THE DISSERTATION

Extension of Cure Rate Model When Cured Is Partially Known

by Yu Wu

Dissertation Director: Yong Lin, Ph.D

When there is evidence of long term survivors, cure rate models have been used by researchers to model the tail behavior of the survival curve. Mixture models were traditionally used, and different parameter estimation approaches for parametric and semi-parametric models have been suggested. A common aspect of the traditional cure rate models is that they implicitly assume there is no additional information about the status of cure, thus the indicator of cure has been modeled as latent variable. This assumption is not entirely valid in many cases, when some diagnostic procedure can provide information about the status of cure. This dissertation proposes a novel extension to incorporate the additional information about status of cure in the cure rate models. It also shows that, with this additional information, more efficient estimator can be obtained. The efficiency gain increases with better sensitivity and specificity of the diagnostic procedure. The efficiency gain is larger when the censoring rate is high.

This extension can be applied when the latency part is modeled parametrically, semi-parametrically, or non-parametrically. Both proportional hazards (PH) cure rate models and accelerated failure time (AFT) cure rate models can use this model extension. Simulation study and a case study results are presented.

Preface

This dissertation is organized as following:

Chapter 1 describes the background of the cure rate model, and presents the motivating example for the dissertation.

Chapter 2 reviews the literature for the existing modeling and estimation methods for the cure rate models.

Chapter 3 develops the proposed new modeling method to include additional diagnostic information into the cure rate models.

Chapter 4 presents a theoretical study of the relative efficiency under special cases. Simulation study is performed using semi-parametric AFT and PH cure rate models and incorporating additional diagnostic information.

Chapter 5 applies the proposed new modeling method to the motivating example.

Conclusions and future works are presented in Chapter 6.

Acknowledgements

First and foremost, I want to thank my advisor, Dr. Yong Lin, for his continuous support throughout the course of this doctoral research journey. He has always been available to listen and to give advice. I started this doctoral program by taking his classes in Advanced Theories. He has been training me since the onset of my doctoral program. He is the one who introduced me to the cure rate modeling literature; he showed me different ways to approach research problems; and he is the one that I turned to whenever I encountered obstacles in my research. It would have been impossible for me to accomplish the goal of this research without his guidance and insight.

I owe enormous gratitude to Dr. Weichung Joe Shih. He provided critical guidance at each key step of this research, from the initiation step, to the stage of the thesis proposal, and leading to this thesis. The range of the guidance involves every aspect from research methods to technical writing. I benefit from them not only for this thesis research topic, but also for future research.

I am indebted to Dr. Shou-En Lu. Dr. Lu worked with me not only for this research project, but also for a previous project regarding survey weighting and non-response. The project taught me ways to conduct research. I also want to express my sincere thanks to Dr. Yujun Wu. Being an adjunct faculty member, he has to balance his work and research but still found time to give me input and advice. I had many discussion sections with Dr. Lu and Dr. Wu. I gained many ideas and insights from them during these thought provoking discussions.

I also want to thank my colleagues and supervisors. Having a fulltime job while working on the Ph.D. program has been a real challenge for me. They supported me with a flexible schedule, and helped me to cover the workload.

Last, but not least, I want to thank my family. To my parents, for helping me with

my education to prepare for this doctoral program, for helping with the housework at their old age; to my beloved wife, who has been handling most of the chores and taking care of the two babies during the past five years; and to my daughter Shannon and my son Ethan, whose wonderful smiles and beautiful artworks always cheer me up.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

A common assumption in survival analysis is that if there is no censoring, the event of interest will eventually occur. Under this assumption, the survival curve will eventually decrease to 0. However, there have been many examples that contradict this assumption. In clinical settings, Wither et al. (1995) showed Kaplan-Meier curves for local control by external beam radiation therapy for squamous carcinoma of tonsil as a function of T-stage and N-stage (Figure 1.1). Patients were followed up for a minimum of 5 years. For each of the K-M curve, a plateau was reached in about 1 to 3 years of follow-up. Taylor (1995) pointed out that because of the kinetics of tumor growth, it is extremely unlikely to have any recurrence later than 5 years after radiation treatment. In sociology setting, Yamaguchi (1992) illustrated several examples of cured subjects: marriage, divorce, birth of second child, career shift, etc. Take career shift as an example, intuitively, after an individual has been in a industry for 10 years, it is unlikely for the individual to shift career to another industry. When the event of interest is not death, the subjects whose event of interest will not occur, are called cured subjects in the literature.

Conventional survival analyses use proportional hazards (PH) model or accelerated failure time (AFT) model. When these models are specified parametrically, the underlying assumption is the event of interest will eventually occur, which is not appropriate for cured subjects. Semi-parametric PH or AFT models can potentially model the cured information by setting the hazard function to 0 after a time threshold. However, for PH model, this time threshold has to be the same for all covariate groups; for AFT model, this time threshold has to be the same after scale change. Furthermore, by

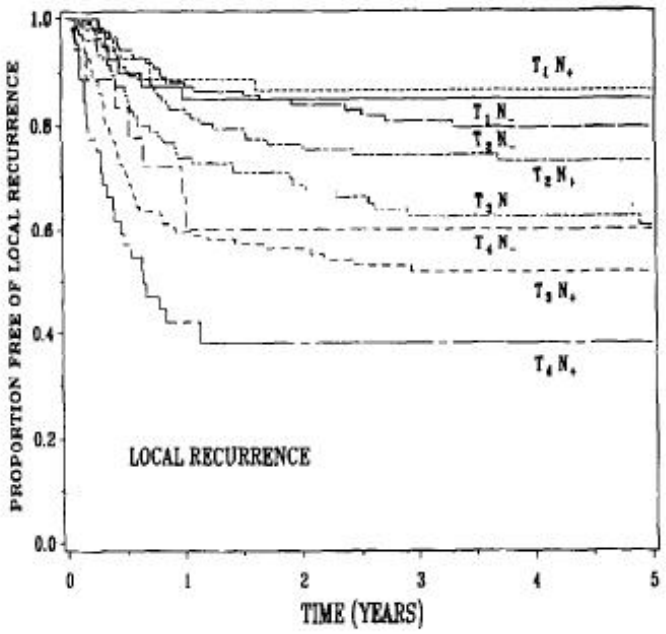


Figure 1.1: KM Curve for local control by external beam radiation therapy for squamous cacinoma of tonsil as a function of TN stage (H.R. Withers et al, 1995)

modeling in this way, long term survivors can not be distinguished from cured subjects.

Alternative modeling method by using cure rate models has been studied by many authors. If we denote T as a non-negative random variable for the failure time, $\mathbf{x}$ and $\mathbf{z}$ as the covariate vectors, $\pi(\mathbf{z})$ as the probability of uncured for a subject, and $f(t|\mathbf{x},\mathbf{z})$ and $S(t|\mathbf{x},\mathbf{z})$ as the density and the survival function for T, respectively. Furthermore, let $f_u(t|\mathbf{x})$ and $S_u(t|\mathbf{x})$ be the density and the survival function for uncured subjects. The basic formulation of cure rate model can be written as the following mixture model:

$$f(t|\mathbf{x},\mathbf{z}) = \pi(\mathbf{z}) f_u(t|\mathbf{x}),$$

or

$$S(t|\mathbf{x},\mathbf{z}) = \pi(\mathbf{z}) S_u(t|\mathbf{x}) + (1 - \pi(\mathbf{z})). \tag{1.1}$$

Different modeling options are available for $\pi(\mathbf{z})$ and $f_u(t|\mathbf{x})$. $\pi(\mathbf{z})$ is called the "incidence" part. It is commonly modeled using logistic regression, although other links or non-linear regression methods could be used. $f_u(t|\mathbf{x})$ is called the "latency"

part. It could be modeled parametrically, semi-parametrically, or non-parametrically. Parametric approaches use different distributions to model the "latency" part. The following distributions have commonly been used: Exponential distribution (Jones et al., 1981; Goldman, 1984; Ghitany et al., 1994); Weibull (Farewell, 1986, 1982); Log-normal (Boag, 1949; Gamel et al., 1990); Gompertz (Gordon,1990a,b; Cantor and Shuster, 1992); Extended generalized gamma (EGG) (Yamaguchi, 1992); and Generalized F (GF) distributions (Peng et al., 1998). The "latency" part can also be modeled nonparametrically or semiparametically, and this leads to semiparametric mixture cure models. Taylor (1995) adopted nonparametric approach to estimating the survival function using Kaplan-Meier estimator without adjusting for the covariates. The popular semiparametric approach is the Cox proportional hazards (PH) model. These include Kuk and Chen (1992), who used Monte Carlo simulation to estimate the parameters in the model; Peng and Dear (2000), who used EM algorithm to estimate the parameters; Sy and Taylor (2000), who used EM algorithm to estimate the parameters. An alternative semiparametric approach uses accelerated failure time (AFT) models. These include Li and Taylor (2002), who used EM algorithm to estimate the parameters; Zhang and Peng (2007), who used EM algorithm and the rank estimator to estimate the parameters. Parametric cure models can achieve the greatest efficiency in estimation if their distributional assumptions are satisfied. However, verifying the assumptions is a challenge in practice. Semiparametric models do not require a distributional assumption. However, it may lose efficiency in estimation as compared to a parametric model when a distribution can be correctly identified.

All the cure rate modeling to date has a common feature: they all assume cured and uncured subjects can not be distinguished in the censored subset. In some cases, this assumption did not take full advantage of the information available in the data, or the innate mechanism of the event. In the radiology example mentioned above, if subjects were unlikely to have events when recurrence did not happen in 5 years of follow-up, we can almost sure that subjects that were followed up more than 5 years without local recurrence event are cured subjects. As another example, consider a study on bone injury in pediatric patients (Leary et al, 2009). One hundred and fifty-seven

children's charts were reviewed to identify the incidence of premature physeal closure (PPC) following physeal fractures of the distal end of the tibia. For 16 children, PPC were identified. Other children were censored. Among the censored children, closure of the growth plate was noted without the events of PPC, thus they can be viewed as cured. Just modeling the data under the usual cure rate model framework would not have taken advantage of this additional cured information.

In many cases, there are diagnostic procedures available to provide further information about whether a subject is cured. It is also worth noting that the additional information about cured status may not be available for all subjects, and all diagnostic procedures are likely to be associated with certain degree of sensitivity and specificity. Complete separation of cured and uncured subjects in the censored subset can hardly be achieved. So, the cure rate model that incorporates the additional information would also need to take into account of the sensitivity and specificity of these additional information. Ideally, when the additional information are of high sensitivity and specificity, more efficient estimating procedure can be constructed.

1.2 Research Questions and Objectives

As the example of pediatric bone injury in the introduction section illustrates, when there are additional sources of information to indicate the cured or uncured status of a patient, current cure rate models can not take advantage of this additional information. The focus of this thesis is how to incorporate this additional information into cure rate models, and produce better estimators.

The objectives of this thesis are:

1. To develop a new model that will incorporate additional diagnostic information into cure rate models.

2. To propose a parameter estimation procedure for this new model.

3. To study the relative efficiency of the new model with and without additional diagnostic procedure information included.

4. To study the asymptotic relative efficiency of the new model theoretically.

The rest of this thesis is organized as follows. In Chapter 2, we will review the literature for the existing modeling and estimation methods. The research methods will be presented in Chapter 3. Theoretical and simulation studies for the proposed model will be presented in Chapter 4. An application of the new methods to a case study of pediatric bone fracture can be found in Chapter 5. Conclusion and future works will be presented in Chapter 6.

Chapter 2

Literature Review

2.1 Likelihood Function for Cure Rate Models

Consider the data in the form $(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i)$, $i = 1, 2, ..., n$, where t_i denotes the observed survival time for the i^{th} patient; δ_i is the censoring indicator with 1 if t_i is uncensored (i.e, observed), and 0 censored; $\mathbf{x}_i$ and $\mathbf{z}_i$ are two covariate vectors. For simplicity, we will use $\mathbf{O} = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i) , i = 1, 2, ..., n\}$ to denote the observed data, and $\boldsymbol{\theta}$ to denote all model parameters. If we model this set of data using cure rate model specified in (1.1), the observed likelihood can be written as:

$$\begin{aligned} L_{obs}(\boldsymbol{\theta}; \mathbf{O}) &= \prod_{i=1}^{n} f(t_i, \delta_i | \mathbf{x}_i, \mathbf{z}_i) = \prod_{i=1}^{n} \left[f(t_i, 1 | \mathbf{x}_i, \mathbf{z}_i) \right]^{\delta_i} \left[f(t_i, 0 | \mathbf{x}_i, \mathbf{z}_i) \right]^{1-\delta_i} \\ &= \prod_{i=1}^{n} \left[\pi(\mathbf{z}_i) f_u(t_i | \mathbf{x}_i) \right]^{\delta_i} \left[\pi(\mathbf{z}_i) S_u(t_i | \mathbf{x}_i) + (1 - \pi(\mathbf{z}_i)) \right]^{1-\delta_i} . \end{aligned} \tag{2.1}$$

Different modeling choices are available for $S_u(t|\mathbf{x})$ and $\pi(\mathbf{z})$. A common modeling choice for the "incidence" part $\pi(\mathbf{z})$, the probability of uncured given covariate vector $\mathbf{z}$, is logistic regression models (for example, Farewell 1982, Peng and Dear 2000). Other link function or non-linear regression may be used. For "latency" part $S_u(t|\mathbf{x})$, early literature have been focused on parametric modeling (either parametric PH or AFT models). In the past decade, there were a lot of interests on semi-parametric PH models, as well as semi-parametric AFT models.

In the following literature review, we will use the data in an ordered form. Denote the distinct event times by $t_{(1)} < t_{(2)} < ... < t_{(w)}$, $w \leq n$. Further let $m_{(i)}$ be the number of events that occurred on $t_{(i)}$, R_i be the risk set right before $t_{(i)}$, $C_{(i)}$ be the censored set in the interval $\left[t_{(i)}, t_{(i+1)}\right)$, and D be the set of observations with events. We will use c_i as the indicator of the status of being not cured for i^{th} patient, i.e., 1 is

not cured (susceptible patient), and 0 is cured (not susceptible patient). c_i is partially observed since $\delta_i = 1$ implies $c_i = 1$.

2.2 Cure Rate Models with Parametric Specification for the Latency Part

Early cure rate model literature had been focused on parametric mixture models. Farewell (1982) used Weibull distribution as the density function for $f_u(t|\mathbf{x})$. If we use k to denote the shape parameter, and use h to denote the scale parameter of Weibull distribution, then we have:

$$f_u(t|\mathbf{x}, h, k) = kh(ht)e^{-(ht)^k}$$

and

$$S_u(t|\mathbf{x}, h, k) = e^{-(ht)^k}. \tag{2.2}$$

For the modeling of the probability $\pi(\mathbf{z})$ of uncured, we can model it using logistic regression:

$$\pi(\mathbf{z}) = \frac{e^{\boldsymbol{\gamma}'\mathbf{z}}}{1 - e^{\boldsymbol{\gamma}'\mathbf{z}}}, \tag{2.3}$$

h can usually be modeled by:

$$h = e^{\boldsymbol{\beta}'\mathbf{x}} h_0. \tag{2.4}$$

Under Weibull assumption, (2.4) may be regarded as both PH and AFT model. Substitute (2.2), (2.3), and (2.4) into (2.1), we can then obtain the MLE of the parameters and asymptotic covariance matrix using Newton-Raphson algorithm.

Other parametric distributional specifications for survival function have been proposed as well. In literature, the following distributions are commonly used: exponential, gamma, Weibull, lognormal, Extended Generalized Gamma (EGG) distribution, and Generalized F (GF) distribution. If the GF distribution is used in parametric cure models, the density function of the failure time distribution of uncured patients is then given by

$$f_u(t|x) = p\lambda(\lambda t)^{ps_1 - 1} \left[1 + (\lambda t)^p\right]^{-s_1 + s_2} B^{-1}(s_1, s_2),$$

where $\lambda = (s_1/s_2)^{1/p}\, e^{-\boldsymbol{\beta}'\mathbf{x}}$ is a scale parameter with positive shape parameters s_1, s_2, p, and $B(s_1, s_2)$ is the beta function evaluated at s_1 and s_2.

The main advantage of using the GF distribution is its flexibility. As shown in Kalbfleisch and Prentice (1980) and Hogg and Ciampi (1985), it includes all the distributions mentioned above as special cases. For example, when $s_1 \to \infty$ or $s_2 \to \infty$, the GF distribution reduces to the EGG distribution with the following density function

$$f_u(t|x) = \begin{cases} \frac{|q|\left(q^{-2}\right)^{q^{-2}} p\lambda(\lambda t)^{p/q-1} \exp\left[-q^{-2}(\lambda t)^{qp}\right]}{\Gamma(q^{-2})}, & q \neq 0 \\ p(\sqrt{2\pi}t)^{-1} \exp(-p^2 \left[\log(\lambda t)\right]^2/2), & q = 0, \end{cases}$$

where $\lambda = e^{-\beta' x}$ and

$$q = \begin{cases} 1/\sqrt{s_1}, & \text{if } s_2 \to \infty \\ -1/\sqrt{s_2}, & \text{if } s_1 \to \infty. \end{cases}$$

The EGG distribution further reduces to

- the lognormal when $q = 0$,
- the Weibull when $q = 1$, and
- the exponential when $q = 0$ and $p = 1$.

Hence, GF distribution is a very flexible family of distributions. But, the extra parameters in the EGG or GF models may cause a loss of efficiency in estimation if the underlying distribution is close to a simpler model. Similar to Weibull mixture model, the MLE of the parameters in the extended Gamma family and the generalized F distribution can also be found through Newton-Raphson algorithm, although the estimation procedure is more complicated because of the need for numerical integration for the survival function.

Use of EGG distribution in AFT models is discussed by Yamaguchi (1992). Use of GF distribution in AFT models is discussed by Peng et al (1998).

2.3 Cure Rate Models with Semi-Parametric PH Specification for the Latency Part

As Peng and Dear (2000) mentioned, "a problem with parametric models is that it is difficult to verify the distributional assumptions used in the models". This has been a general problem for parametric models in survival data analysis. To verify the distributional assumptions used in the cure rate models is even more challenging because of the mixture nature of the model. For this reason, semi-parametric models are proposed, to relax the parametric assumption, at the same time still maintain the possibility to include model covariates. Cure rate models with semi-parametric PH specification for latency part is one of the semi-parametric models proposed. PH model is a popular survival analysis model, and allows one to interpret the covariates as proportionately increase or decrease the hazard of failure at any time.

The key issue for cure rate models with semi-parametric PH specification is the parameter estimation. In the next few sub-sections, we will look at why this issue exists, and some solutions that have been proposed in the literature.

2.3.1 Difficulty with Partial Likelihood Approach

Under proportional hazards assumption without mixture models, hazard function is specified as $h(t|\mathbf{x}) = e^{\boldsymbol{\beta}'\mathbf{x}} h_0(t)$, where $h_0(t)$ is the baseline hazard function when $\mathbf{x} = \mathbf{0}$ that can take any non-negative functional form. In Cox's original paper (1972, 1975), he used a partial likelihood approach to cancel out $h_0(t)$. The partial likelihood can be computed as:

$$L_{cond}(\boldsymbol{\theta}; \mathbf{x}) = \prod_{i=1}^{w} \frac{e^{\boldsymbol{\beta}'\mathbf{x}_i} h_0(t)}{\sum\limits_{l \in R_i} e^{\boldsymbol{\beta}'\mathbf{x}_l} h_0(t)} = \prod_{i=1}^{w} \frac{e^{\boldsymbol{\beta}'\mathbf{x}_i}}{\sum\limits_{l \in R_i} e^{\boldsymbol{\beta}'\mathbf{x}_l}}.$$

Note that $L_{cond}(\theta)$ is a function of $\boldsymbol{\beta}$, and does not involve $h_0(t)$. As a result, estimation for $\boldsymbol{\beta}$ can be carried out using Newton-Raphson algorithm. After $\boldsymbol{\beta}$ is estimated, following Breslow (1972, discussion on Cox's paper), we can estimate the baseline hazard function as a step function:

$$\hat{h}_0(t) = \hat{h}_i \text{ for } t_{(i-1)} < t \leq t_{(i)},$$

where

$$\hat{h}_i = \frac{m_{(i)}}{\left(t_{(i)} - t_{(i-1)}\right) \sum\limits_{l \in R_i} e^{\boldsymbol{\beta}' \mathbf{x}_l}}.$$

Then the survival function can be estimated in a similar form as the Nelson-Aalen estimator:

$$\hat{S}_0(t) = \exp\left(-\int_0^t \hat{h}_0(t)dt\right) = \exp\left(-\sum_{i:t_{(i)}<t} \frac{m_{(i)}}{\sum\limits_{l \in R_i} e^{\boldsymbol{\beta}' \mathbf{x}_l}}\right). \tag{2.5}$$

Alternatively, one may also estimate the survival function in a similar form as the Kaplan-Meier estimate:

$$\hat{S}_0(t) = \prod_{i:t_{(i)}<t} \left(1 - \frac{m_{(i)}}{\sum\limits_{l \in R_i} e^{\boldsymbol{\beta}' \mathbf{x}_l}}\right). \tag{2.6}$$

When there is a model mixture, as is in the case of cure rate model, hazard function can be written as:

$$\begin{aligned} h(t|\mathbf{x}, \mathbf{z}) &= \frac{f(t|\mathbf{x}, \mathbf{z})}{S(t|\mathbf{x}, \mathbf{z})} \\ &= \frac{\pi(\mathbf{z}) f_u(t|\mathbf{x})}{\pi(\mathbf{z}) S_u(t|\mathbf{x}) + (1 - \pi(\mathbf{z}))} \\ &= \frac{\pi(\mathbf{z}) S_u(t|\mathbf{x})}{\pi(\mathbf{z}) S_u(t|\mathbf{x}) + (1 - \pi(\mathbf{z}))} h_u(t|\mathbf{x}) \\ &= \frac{\pi(\mathbf{z}) S_u(t|\mathbf{x})}{\pi(\mathbf{z}) S_u(t|\mathbf{x}) + (1 - \pi(\mathbf{z}))} e^{\boldsymbol{\beta}' \mathbf{x}} h_0(t). \end{aligned}$$

Now the ratio between the hazard function and the baseline hazard depends on $S_u(t|\mathbf{x})$. As a result, joint estimation for the baseline hazard function and parameter $\boldsymbol{\beta}$ is needed. The estimation methods described above can not be directly applied to cure rate models.

2.3.2 Marginal Likelihood Approach

To solve the difficulty of using partial likelihood, Kuk and Chen (1992) proposed the use of the marginal likelihood for the parameter estimation in the cure rate models with the semi-parametric PH specification. In their method, the following likelihood

function was used:

$$L_m(\boldsymbol{\theta};\mathbf{O}) = \int_{t_{(1)}<\ldots<t_{(k)}} \cdots \int \prod_{i=1}^{w} \left\{ \pi(\mathbf{z}_{(i)}) f_u(t_{(i)}|\mathbf{x}_{(i)}) \prod_{l\in C_i} \left[1-\pi(\mathbf{z}_{(l)}) + \pi(\mathbf{z}_{(l)}) S_u(t_{(i)}|\mathbf{x}_{(l)})\right] \right\} dt_{(1)}\ldots dt_{(m)}.$$

This integration sums up all the likelihoods under the restriction of $t_{(1)} < t_{(2)} < \ldots < t_{(k)}$, thus integrates out the time dependent hazard function. They further show that this likelihood function can be written as:

$$L_m(\boldsymbol{\theta};\mathbf{O}) = \sum_{\mathbf{c}\in\boldsymbol{\Omega}_{n-k}} L_m(\boldsymbol{\theta};\mathbf{O}|\mathbf{c}), \tag{2.7}$$

where $\boldsymbol{\Omega}_{n-w}$ is the collection of $(n-w)$-tuple of 0 and 1, and

$$L_m(\boldsymbol{\theta};\mathbf{O}|\mathbf{c}) = \prod_{i\in D} \pi(\mathbf{z}_i) \prod_{i=1}^{n-w} \left[\pi(\mathbf{z}_i)^{c_i}(1-\pi(\mathbf{z}_i))^{1-c_i}\right] \prod_{i=1}^{w} \frac{e^{\boldsymbol{\beta}'\mathbf{x}_{(i)}}}{\sum_{j=i}^{m} e^{\boldsymbol{\beta}'\mathbf{x}_{(j)}} + \sum_{l\in C_{(j)}} c_l e^{\boldsymbol{\beta}'\mathbf{x}_l}}.$$

For censored observations, we can not determine whether they are censored because of cure. By computing $L_m(\boldsymbol{\theta};\mathbf{O})$ as the sum of $L_m(\boldsymbol{\theta};\mathbf{O}|\mathbf{c})$, this approach considers all the possibilities of cured/non-cured of censored observations, then sums over all the possibilities. One issue with this approach is that, even with modest number of censored observations, the number of observations in $\boldsymbol{\Omega}_{n-w}$ (which is 2^{n-w}) may be too large for an exact computation. They used Monte Carlo simulation to draw samples from $\boldsymbol{\Omega}_{n-w}$, thus provide an estimation for $L_m(\boldsymbol{\theta};\mathbf{O})$. Estimation for $\boldsymbol{\theta}$ is obtained by maximizing $L_m(\boldsymbol{\theta};\mathbf{O})$.

Variance for $\boldsymbol{\theta}$ is obtained by inverting the information matrix, plus the sampling variance due to Monte Carlo simulation.

2.3.3 EM Algorithm

Normally, when the sample size is large, better estimates can be obtained. In the marginal likelihood approach, because of the computational limitation, when the sample size is large, still a relatively small number of Monte Carlo samples can be taken from $\boldsymbol{\Omega}_{n-w}$. So, the full advantage of large sample size is not taken. Furthermore, because the marginal likelihood approach used a very different approach from traditional survival analysis, it is hard to take advantage from existing software package. These

issues can potentially be addressed by the use of EM algorithm, as proposed by Sy and Taylor (2000), Peng and Dear (2000) separately.

By introducing partially observed c_i into the likelihood, we can write the complete likelihood as:

$$\begin{aligned} L_C(\boldsymbol{\theta}; \mathbf{O}, \mathbf{c}) &= L_C(\boldsymbol{\beta}, \boldsymbol{\gamma}, S_0; \mathbf{O}, \mathbf{c}) \\ &= \prod_{i=1}^{n} \left[\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i)\right]^{\delta_i} \left[\pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i)\right]^{c_i(1-\delta_i)} \left[1-\pi(\mathbf{z}_i)\right]^{(1-c_i)(1-\delta_i)}, \quad (2.8) \end{aligned}$$

where

$$\delta_i = 1 \text{ if event is observed, } \delta_i = 0 \text{ if censored,}$$

and

$$c_i = 1 \text{ if uncured observation, } c_i = 0 \text{ if cured.}$$

The EM algorithm can then be carried out using the following steps.

Initial value: The EM algorithm starts with an initial value $\boldsymbol{\theta}^{(0)}$.

E-step: In the E-step of $(r+1)^{th}$ iteration, the expected value of the log-likelihood can be separated into two parts:

$$\begin{aligned} &Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(r)}, \mathbf{O}) \\ &= E\left[\log(L_C)|\mathbf{O}, \boldsymbol{\theta}^{(r)}\right] \\ &= \sum_{i=1}^{n} \left[g_i^{(r)} \log(\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)})) + (1-g_i^{(r)}) \log(1-\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)}))\right] \\ &\quad + \sum_{i=1}^{n} g_i^{(r)} \log(S_u^{(r)}(t_i|\mathbf{x}_i, \boldsymbol{\beta}^{(r)}) + \delta_i \log(h_u^{(r)}(t_i|\mathbf{x}_i, \boldsymbol{\beta}^{(r)}), \end{aligned}$$

where

$$\begin{aligned} g_i^{(r)} &= E\left(c_i|\mathbf{O}, \boldsymbol{\theta}^{(r)}\right) \\ &= \delta_i + (1-\delta_i) \frac{\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)}) S_u^{(r)}(t_i|\mathbf{x}_i, \boldsymbol{\beta}^{(r)})}{1-\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)}) + \pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)}) S_u^{(r)}(t_i|\mathbf{x}_i, \boldsymbol{\beta}^{(r)})}. \end{aligned}$$

M-step: M-step in the $(r+1)^{th}$ iteration maximizes the expected complete log-likelihood function with respect to θ to obtain $\theta^{(r+1)}$, where the expected complete

log-likelihood function is the sum of the following two functions:

$$l_1(\boldsymbol{\gamma}|g_i^{(r)}, \mathbf{O}) = \sum_{i=1}^{n} \left[g_i^{(r)} \log(\pi(\mathbf{z}_i)) + (1 - g_i^{(r)}) \log(1 - \pi(\mathbf{z}_i)) \right], \tag{2.9}$$

$$l_2(\boldsymbol{\beta}|g_i^{(r)}, \mathbf{O}) = \sum_{i:d_i=0} \left[\delta_i \log(h_u(t_i|\mathbf{x}_i)) + g_i^{(r)} \log(S_u(t_i|\mathbf{x}_i)) \right]. \tag{2.10}$$

When a logistic regression model is used to model $\pi(z_i)$, equation (2.9) is the log-likelihood function of a logistic regression model for values arising from a binomial distribution with the response probability $\pi_i = \exp(\boldsymbol{\gamma}'\mathbf{z}_i)/(1 + \exp(\boldsymbol{\gamma}'\mathbf{z}_i))$. The first part $l_1(\boldsymbol{\gamma}|g_i^{(r)}, \mathbf{O})$ can be maximized by optimization methods such as the Newton-Raphson method. This maximization procedure can be carried out in most standard logistic regression program to obtain $\boldsymbol{\gamma}^{(r+1)}$. Other link functions can similarly be handled.

For equation (2.10), $\boldsymbol{\beta}^{(r+1)}$ can be obtained by a program for the Cox proportional hazards model that accepts covariates with fixed coefficients (Peng, 2003a). After getting $\boldsymbol{\beta}^{(r+1)}$, following the similar method as in (2.5), $S_0^{(r+1)}$ can be obtained by:

$$S_0^{(r+1)}(t) = \exp\left(- \sum_{i:t_{(i)}<t} \frac{m_{(i)}}{\sum_{l \in R_i} g_i^{(r)} e^{\boldsymbol{\beta}^{(r+1)'}\mathbf{x}_l}} \right). \tag{2.11}$$

To avoid identifiability issue, censored observations that are later than the largest event times are considered as non-susceptible observations (referred to as 0-tail completion for survival function estimate).

Iteration: The algorithm is iterated until $\left\| \theta^{(r+1)} - \theta^{(r)} \right\|$ is sufficiently small.

For variance estimations, Sy and Taylor (2000) computed the observed full information matrix, and inverted this matrix to get an approximations of the asymptotic variance. Because of the large number of parameters involved in baseline hazard function, this is potentially a large matrix to inverse. Numerical instability may also occur during the inversion. Alternatively, Peng (2003a) proposed the use of bootstrapping methods for variance estimations.

2.4 Cure Rate Models with Semi-Parametric AFT Specification for the Latency Part

AFT model allows one to interpret the effect of a covariate as either decelerating or accelerating the timing of the event (Li and Taylor, 2002). Li and Taylor (2002) further gave an example in cancer studies, when the event is related to tumor development, it allows a possible and quite appealing interpretation in terms of the kinetics of growth of tumor cells. Several authors explored the use of semi-parametric AFT specification for the survival function in cure rate models. For AFT model, the latency portion is modeled as:

$$\log(T) = \boldsymbol{\beta}^{'}\mathbf{x} + \epsilon, \tag{2.12}$$

where ϵ is an error term with the density function f_0 and the survival function S_0. Given $\boldsymbol{\beta}$ and $\mathbf{x}$, the conditional survival function of T can be derived as $S_0(\log(t)-\boldsymbol{\beta}^{'}\mathbf{x})$. When the distribution of ϵ is modeled non-parametrically, (2.12) is a semi-parametric AFT model.

Li and Taylor (2002), Zhang and Peng (2007) separately presented the formulation and parameter estimation when the survival function is specified using semi-parametric AFT model. Similar to the EM approach for PH model, under AFT assumption, if we introduce partially observed c_i into the likelihood, we can write the complete likelihood as:

$$\begin{aligned} L_C(\boldsymbol{\theta};\mathbf{O},\mathbf{c}) &= L_C(\boldsymbol{\beta},\boldsymbol{\gamma},S_0;\mathbf{O},\mathbf{c}) \\ &= \prod_{i=1}^{n}\left[\pi(\mathbf{z}_i)\frac{f_0(\log(t_i)-\boldsymbol{\beta}^{'}\mathbf{x}_i)}{t_i}\right]^{\delta_i}\left[\pi(\mathbf{z}_i)S_0(\log(t_i)-\boldsymbol{\beta}^{'}\mathbf{x}_i)\right]^{c_i(1-\delta_i)} \\ &\qquad \times\left[1-\pi(\mathbf{z}_i)\right]^{(1-c_i)(1-\delta_i)}, \end{aligned} \tag{2.13}$$

where

$$\delta_i = 1 \text{ if event is observed, } \delta_i = 0 \text{ if censored,}$$

and

$$c_i = 1 \text{ if uncured observation, } c_i = 0 \text{ if cured.}$$

The EM algorithm can then be carried out using the following steps.

Initial value: The EM algorithm starts with an initial value $\boldsymbol{\theta}^{(0)}$.

E-step: In the E-step of $(r+1)^{th}$ iteration, the expected value of the log-likelihood can be separated into two parts, plus a constant term:

$$\begin{aligned}
&Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(r)}, \mathbf{O}) \\
&= E\left[\log(L_C)|\mathbf{O}, \boldsymbol{\theta}^{(r)}\right] \\
&= \sum_{i=1}^{n}\left[g_i^{(r)}\log(\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)})) + (1-g_i^{(r)})\log(1-\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)}))\right] \\
&\quad + \sum_{i=1}^{n} g_i^{(r)}\log(S_0(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i|\boldsymbol{\beta}^{(r)}) + \delta_i\log(h_0(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i|\boldsymbol{\beta}^{(r)}) \\
&\quad - \sum_{i=1}^{n}\delta_i\log(t_i),
\end{aligned}$$

where

$$\begin{aligned}
g_i^{(r)} &= E\left(c_i|\mathbf{O}, \boldsymbol{\theta}^{(r)}\right) \\
&= \delta_i + (1-\delta_i)\frac{\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)})S_0^{(r)}(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i|\boldsymbol{\beta}^{(r)})}{1-\pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)}) + \pi(\mathbf{z}_i|\boldsymbol{\gamma}^{(r)})S_0^{(r)}(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i|\boldsymbol{\beta}^{(r)})}
\end{aligned}$$

M-step: M-step in the $(r+1)^{th}$ iteration maximizes the expected complete log-likelihood function with respect to θ to obtain $\theta^{(r+1)}$, where the expected complete log-likelihood function is the sum of the following two functions:

$$l_1(\boldsymbol{\gamma}|g_i^{(r)}, \mathbf{O}) = \sum_{i=1}^{n}\left[g_i^{(r)}\log(\pi(\mathbf{z}_i)) + (1-g_i^{(r)})\log(1-\pi(\mathbf{z}_i))\right] \tag{2.14}$$

$$l_2(\boldsymbol{\beta}|g_i^{(r)}, \mathbf{O}) = \sum_{i=1}^{n}\left[g_i^{(r)}\log(S_0(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i) + \delta_i\log(h_0(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i)\right] \tag{2.15}$$

When a logistic regression model is used to model $\pi(z_i)$, equation (2.14) is the log-likelihood function of a logistic regression model for values arising from a binomial distribution with the response probability $\pi_i = \exp(\boldsymbol{\gamma}'\mathbf{z}_i)/(1+\exp(\boldsymbol{\gamma}'\mathbf{z}_i))$. The first part $l_1(\boldsymbol{\gamma}|g_i^{(r)}, \mathbf{O})$ can be maximized by optimization methods such as the Newton-Raphson method. This maximization procedure can be carried out in most standard logistic regression programs to obtain $\boldsymbol{\gamma}^{(r+1)}$. Other link functions can similarly be handled.

For equation (2.15), maximization of $l_2(\boldsymbol{\beta}|g_i^{(r)}, \mathbf{O})$ involves the joint estimation of $\boldsymbol{\beta}$ and S_0. Li and Taylor (2002) used a method based on Ritov (1982) to estimate $\boldsymbol{\beta}$. By

setting $\frac{\partial l_2(\boldsymbol{\beta}|g_i^{(r)},\mathbf{O})}{\partial \boldsymbol{\beta}} = 0$, the resulting estimating equations are given by:

$$\sum_{i=1}^{n} \mathbf{x}_i \left[-\delta_i \frac{f_0'}{f_0}(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i) + g_i^{(r)}\,(1-\delta_i)\,\frac{f_0}{S_0}(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i) \right] = 0 \tag{2.16}$$

Following Ritov (1982), we replace $-\frac{f_0'}{f_0}$ by a reasonable score function g. Equation 2.16 is updated to:

$$\sum_{i=1}^{n} \mathbf{x}_i \left[\delta_i g(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i) + g_i^{(r)}\,(1-\delta_i)\,\frac{\int_{\log(t_i)-\boldsymbol{\beta}'\mathbf{x}_i}^{\infty} g(u)dF_0(u)}{S_0(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i)} \right] = 0. \tag{2.17}$$

Ritov further suggested centering $\mathbf{x}_i$ to account for the unknown intercept term, and use the equation:

$$\sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}}) \left[\delta_i g(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i) + g_i^{(r)}\,(1-\delta_i)\,\frac{\int_{\log(t_i)-\boldsymbol{\beta}'\mathbf{x}_i}^{\infty} g(u)dF_0(u)}{S_0(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i)} \right] = 0 \tag{2.18}$$

Because S_0 is unknown, Li and Taylor suggested to replace it by the estimate of $S_0^{\boldsymbol{\beta}}$, where $S_0^{\boldsymbol{\beta}}$ is formed by the Kaplan-Meier estimator using the residuals $v_i = t_i \exp\left(-\boldsymbol{\beta}'\mathbf{x}_i\right)$, as

$$S_0^{\beta}(v) = \prod_{i:v_{(i)}<v} \left(1 - \frac{m_{(i)}}{m_{(i)} + \sum_{l\in R_i} g_l^{(r)}} \right) \tag{2.19}$$

To avoid identifiability issue, zero-tail completion was used for survival function estimate.

Finally, the final estimation equations are written as:

$$\begin{aligned} \Psi(\boldsymbol{\beta};g) &= \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}}) \left[\delta_i g(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i) + g_i^{(r)}\,(1-\delta_i)\,\frac{\int_{\log(t_i)-\boldsymbol{\beta}'\mathbf{x}_i}^{\infty} g(u)dF_0^{\boldsymbol{\beta}}(u)}{S_0^{\boldsymbol{\beta}}(\log(t_i) - \boldsymbol{\beta}'\mathbf{x}_i)} \right] \\ &= 0. \end{aligned} \tag{2.20}$$

Since $\Psi(\boldsymbol{\beta};g)$ is not a monotonic function of $\boldsymbol{\beta}$, Li and Taylor suggested to use grid search to obtain the estimates of $\boldsymbol{\beta}$ and S_0.

Zhang and Peng (2007) commented on a few issues of Li and Taylor's methods. First, because the maximization of $\Psi(\boldsymbol{\beta};g)$ requires grid search, the success of this

maximization step may depend on the initial values, and may fail to produce a consistent estimate. Second, the choice of $g(\bullet)$ function is arbitrary. They proposed to use a rank estimation procedure based on Wei (1992) to estimate $\boldsymbol{\beta}$. In their estimating procedure, because

$$\delta_i \log(g_i^{(r)}) = 0,$$

and

$$\delta_i g_i^{(r)} = \delta_i,$$

they rewrote (2.15) as

$$l_2(\boldsymbol{\beta}|g_i^{(r)}, \mathbf{O}) = \sum_{i=1}^{n} \left[g_i^{(r)} \log\left(S_0(\log(t_i) - \boldsymbol{\beta}^{'}\mathbf{x}_i \right) + \delta_i \log\left(g_i^{(r)} h_0(\log(t_i) - \boldsymbol{\beta}^{'}\mathbf{x}_i) \right) \right] \tag{2.21}$$

This can be taken as the log likelihood of the AFT model with

$$\log(T) = \boldsymbol{\beta}^{'}\mathbf{x} + \epsilon^*,$$

where the hazard function of ϵ^* is

$$h(\epsilon^*) = g_i^{(r)} h_0(\epsilon).$$

The rank estimation procedure considers the following PH model with fixed regression coefficient $\boldsymbol{\varpi} = \mathbf{0}$:

$$h_{ph}(\epsilon^*) = g_i^{(r)} h_0(\epsilon) \exp(\boldsymbol{\varpi}^{'}\mathbf{x}). \tag{2.22}$$

Pretending that we do not know the true value of $\boldsymbol{\varpi}$, we take the derivative of the logarithm of the partial likelihood function for model (2.22) with respect to $\boldsymbol{\varpi}$**:**

$$\Psi(\boldsymbol{\varpi}) = \sum_{i=1}^{n} \delta_i \left(\mathbf{x}_i - \frac{\sum_{j=1}^{n} \mathbf{x}_j g_j^{(r)} e^{\boldsymbol{\varpi}^{'}\mathbf{x}_j} I\left(\epsilon_j^* \geq \epsilon_i^*\right)}{\sum_{j=1}^{n} g_j^{(r)} e^{\boldsymbol{\varpi}^{'}\mathbf{x}_j} I\left(\epsilon_j^* \geq \epsilon_i^*\right)} \right).$$

Plug in $\boldsymbol{\varpi} = \mathbf{0}$, this becomes

$$\Psi(\mathbf{0}) = \sum_{i=1}^{n} \delta_i \left(\mathbf{x}_i - \frac{\sum_{j=1}^{n} \mathbf{x}_j g_j^{(r)} I\left(\epsilon_j^* \geq \epsilon_i^*\right)}{\sum_{j=1}^{n} g_j^{(r)} I\left(\epsilon_j^* \geq \epsilon_i^*\right)} \right). \tag{2.23}$$

(2.23) is a function of $\boldsymbol{\beta}$. It can be further extended to include a weight function:

$$\Psi(\boldsymbol{\beta};k) = \sum_{i=1}^{n} \delta_i k\left(\epsilon_i^*\right) \left(\mathbf{x}_i - \frac{\sum_{j=1}^{n} \mathbf{x}_j g_j^{(r)} I\left(\epsilon_j^* \geq \epsilon_i^*\right)}{\sum_{j=1}^{n} g_j^{(r)} I\left(\epsilon_j^* \geq \epsilon_i^*\right)} \right). \tag{2.24}$$

Using a Gehan type weight function

$$k(u) = \frac{1}{n}\sum_{j=1}^{n} I\left(\epsilon_j^* \geq u\right) g_j^{(r)},$$

the estimation function (2.24) can be simplified as

$$\Psi(\boldsymbol{\beta};k) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \left(\mathbf{x}_i - \mathbf{x}_j\right) g_j^{(r)} I\left(\epsilon_j^* \geq \epsilon_i^*\right). \tag{2.25}$$

This estimating equation $\Psi(\boldsymbol{\beta};k)$ can be taken as the gradient of the convex function

$$L_G(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \left|\epsilon_j^* - \epsilon_i^*\right| g_j^{(r)} I\left(\epsilon_j^* \geq \epsilon_i^*\right). \tag{2.26}$$

Finding the root of $\Psi(\boldsymbol{\beta};k) = 0$ is equivalent to minimizing $L_G(\boldsymbol{\beta})$. $L_G(\boldsymbol{\beta})$ can be minimized by the linear programming method.

After $\boldsymbol{\beta}^{(r+1)}$ is estimated, $S_0^{(r+1)}$ can be estimated by the residuals $v_i = t_i \exp\left(-\boldsymbol{\beta}^{(r+1)'}\mathbf{x}_i\right)$ as

$$S_0^{(r+1)}(v) = \exp\left(-\sum_{i:v_{(i)}<v} \frac{m_{(i)}}{\sum\limits_{l\in R_i} g_i^{(r)}}\right). \tag{2.27}$$

To avoid identifiability issue, zero-tail completion was used for survival function estimate.

Iteration: The algorithm is iterated until $\left\|\theta^{(r+1)} - \theta^{(r)}\right\|$ is sufficiently small.

For variance estimation, both Li and Taylor (2002) and Zhang and Peng (2007) used the bootstrapping methods.

2.5 Cure Rate Models with Non-Parametric Specification for the Latency Part

Non-parametric specification for the latency part was considered by Taylor (1995). The estimation follows the same line as the EM algorithm for the cure rate models with PH specification for the latency part, with all $\boldsymbol{\beta}$ replaced by 0 in the derivation. Survival function for the latency part was estimated by a Kaplan-Meier type estimator as:

$$S_u(t) = \prod_{i:t_{(i)}<t}\left(1 - \frac{m_{(i)}}{m_{(i)} + \sum\limits_{l\in R_i} g_l^{(r)}}\right) \tag{2.28}$$

Same as other estimating scenarios, zero-tail completion was used for survival function estimate.

For variance estimation, Taylor suggested the use of the observed information matrix, but noted the potential issue of a large number of nuisance parameters in the information matrix.

Chapter 3

Cure Rate Models with Sensitivity and Specificity

3.1 Model Specification

Suppose in addition to $(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i)$, for censored patients, we also observed the result d_i from a diagnostic procedure, where d_i is 1 if i^{th} patient is diagnosed as cured, 0 if 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 identified as having the condition). Specificity measures the proportion of negatives which are correctly identified (e.g. the percentage of healthy people who are identified as not having the condition). Assume that given censored, i.e., $\delta_i = 0$, d_i is independent of t_i and the diagnostic procedure has a sensitivity of p_0 and a specificity of $1-p_1$. For a validated diagnostic procedure, we will have $p_0 \geq p_1$. Notice that the sensitivity and the specificity here are the corresponding conditional probabilities given censored, they might be different from the sensitivity and the specificity from the population. We might model p_0 and p_1, but for simplicity, we assume that the conditional probabilities do not depend on any covariates. Denote $\mathbf{O} = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i, d_i), i = 1, 2, ..., n\}$. If the results from the diagnostic procedure are available for all subjects, the observed likelihood becomes:

$$L_{obs}(\boldsymbol{\theta}; \mathbf{O}) = \prod_{i=1}^{n} [\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i)]^{\delta_i} \Big[{p_1}^{d_i} (1-p_1)^{1-d_i} \pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i) + p_0^{d_i} (1-p_0)^{1-d_i} (1-\pi(\mathbf{z}_i)) \Big]^{1-\delta_i}, \tag{3.1}$$

because for uncensored patients ($\delta_i = 1$), the contribution to the likelihood is the same as that in (2.1); while for censored patients ($\delta_i = 0$), with the independence assumption of d_i and t_i, the contribution is $p_1^{d_i}(1-p_1)^{1-d_i}\pi_i S_u(t_i|\mathbf{x}_i)$ if they are uncured, and the contribution is $p_0^{d_i}(1-p_0)^{1-d_i}(1-\pi_i)$ if they are cured.

If the diagnostic procedure results are not available for all the censored subjects, let $\eta_i = 1$ denote the result of diagnostic procedure is available for a subject, and $\eta_i = 0$ to denote the result is not available. Let $\mathbf{O} = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i, \eta_i, d_i), i = 1, 2, ..., n\}$, the observed likelihood can be written as:

$$\begin{aligned} L_{obs}(\boldsymbol{\theta}; \mathbf{O}) = \prod_{i=1}^{n} & \left[\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i)\right]^{\delta_i} \\ & \times \left[{p_1}^{d_i}(1-p_1)^{1-d_i}\pi(\mathbf{z}_i)S_u(t_i|\mathbf{x}_i) + p_0^{d_i}(1-p_0)^{1-d_i}\left(1-\pi(\mathbf{z}_i)\right)\right]^{(1-\delta_i)\eta_i} \\ & \times \left[\pi(\mathbf{z}_i)S_u(t_i|\mathbf{x}_i) + (1-\pi(\mathbf{z}_i))\right]^{(1-\delta_i)(1-\eta_i)}. \end{aligned} \tag{3.2}$$

Notice that when $p_0 = p_1$, (3.2) reduces to (2.1) except a constant multiplier, meaning that if both sensitivity and $(1 -$ specificity$)$ are the same, then with and without the diagnostic information are the same.

As in the literature, logistic regression or other link functions or nonlinear regression can be used to model the "incidence" part $\pi(\mathbf{z})$ of the mixture model. Parametric, semiparametric (PH or AFT), or nonparametric methods can be used to model the "latency" part $S_u(t|\mathbf{x})$ of the mixture model. Direct estimation of model parameters in (3.2) now is even harder than that for the traditional mixture cure model in (2.1) since we have two more additional parameters p_0 and p_1 to estimate. The good news is that the EM algorithm is still a good choice for the parameter estimations. The following section provides the details of the EM procedure.

3.2 Estimation for the Proposed Model Using EM Algorithm

For the parameter estimation of the mixture cure model with additional information in the previous section, we follow an EM algorithm similar to Peng (2003a). Denote the observed data as $\mathbf{O} = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i, \eta_i, d_i), i = 1, 2, ..., n\}$. If we introduce c_i as the indicator of the status of being not cured for i^{th} patient, i.e., 1 is not cured (susceptible patient), and 0 is cured. c_i is partially observed since $\delta_i = 1$ implies $c_i = 1$. Then we have:

$$d_i|\,(c_i = 0, \delta_i = 0) \sim Bernoulli(p_0),$$

and

$$d_i|\,(c_i = 1, \delta_i = 0) \sim Bernoulli(p_1).$$

With the introduction of c_i, let $f(t_i, d_i, c_i, \eta_i, \delta_i|\mathbf{x}_i, \mathbf{z}_i)$ be the joint pdf of $(t_i, d_i, c_i, \eta_i, \delta_i)$ given covariates $\mathbf{x}_i$ and $\mathbf{z}_i$, the complete log-likelihood can be written as:

$$\begin{aligned}
l(\theta; \mathbf{O}, \mathbf{c}) &= \log L_c(\boldsymbol{\theta}; \mathbf{O}, \mathbf{c}) \\
&= \log \prod_{i=1}^{n} f(t_i, d_i, c_i, \eta_i, \delta_i|\mathbf{x}_i, \mathbf{z}_i) \\
&= \log \prod_{i=1}^{n} \left[f(t_i, d_i, c_i, \eta_i, 1|\mathbf{x}_i, \mathbf{z}_i)\right]^{\delta_i} \left[f(t_i, d_i, c_i, \eta_i, 0|\mathbf{x}_i, \mathbf{z}_i)\right]^{1-\delta_i} \\
&= \log \prod_{i=1}^{n} \left[f(t_i, d_i, c_i, 0, 1|\mathbf{x}_i, \mathbf{z}_i)\right]^{\delta_i} \\
&\qquad \times \left[\left\{f(t_i, d_i, 1, \eta_i, 0|\mathbf{x}_i, \mathbf{z}_i)\right\}^{c_i} \left\{f(t_i, d_i, 0, \eta_i, 0|\mathbf{x}_i, \mathbf{z}_i)\right\}^{1-c_i}\right]^{1-\delta_i} \\
&= \log \prod_{i=1}^{n} \left[f(t_i, d_i, c_i, 0, 1|\mathbf{x}_i, \mathbf{z}_i)\right]^{\delta_i} \\
&\qquad \times \left[\left\{f(t_i, d_i, 1, 0, 0|\mathbf{x}_i, \mathbf{z}_i)\right\}^{c_i} \left\{f(t_i, d_i, 0, 0, 0|\mathbf{x}_i, \mathbf{z}_i)\right\}^{1-c_i}\right]^{(1-\delta_i)(1-\eta_i)} \\
&\qquad \times \left[\left\{f(t_i, d_i, 1, 1, 0|\mathbf{x}_i, \mathbf{z}_i)\right\}^{c_i} \left\{f(t_i, d_i, 0, 1, 0|\mathbf{x}_i, \mathbf{z}_i)\right\}^{1-c_i}\right]^{(1-\delta_i)\eta_i} \\
&= \log \prod_{i=1}^{n} \left[\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i)\right]^{c_i\delta_i} \left[\left\{\pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i)\right\}^{c_i} \left(1 - \pi(\mathbf{z}_i)\right)^{1-c_i}\right]^{(1-\delta_i)(1-\eta_i)} \\
&\qquad \times \left[\left\{{p_1}^{d_i}(1-p_1)^{1-d_i}\pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i)\right\}^{c_i} \left\{p_0^{d_i}(1-p_0)^{1-d_i}\left(1 - \pi(\mathbf{z}_i)\right)\right\}^{(1-c_i)}\right]^{(1-\delta_i)\eta_i} \\
&= \sum_{i=1}^{n} \left[c_i \log(\pi(\mathbf{z}_i)) + (1 - c_i)(1 - \delta_i) \log(1 - \pi(\mathbf{z}_i))\right] \\
&\qquad + \sum_{i=1}^{n} \left[c_i\delta_i \log(f_u(t_i|\mathbf{x}_i)) + c_i(1 - \delta_i) \log(S_u(t_i|\mathbf{x}_i))\right] \\
&\qquad + \sum_{i=1}^{n} \left[d_i \log p_1 + (1 - d_i) \log(1 - p_1)\right] c_i(1 - \delta_i)\eta_i \\
&\qquad + \sum_{i=1}^{n} \left[d_i \log p_0 + (1 - d_i) \log(1 - p_0)\right] (1 - c_i)(1 - \delta_i)\eta_i. \qquad (3.3)
\end{aligned}$$

Notice that $(1-c_i)(1-\delta_i) = 1-c_i$ and $c_i\delta_i = \delta_i$. (3.3) can be further simplified to:

$$\begin{aligned} l(\theta; \mathbf{O}, \mathbf{c}) = &\sum_{i=1}^{n} \left[c_i \log(\pi(\mathbf{z}_i)) + (1-c_i) log(1-\pi(\mathbf{z}_i))\right] \\ &+ \sum_{i=1}^{n} \left[\delta_i \log(h_u(t_i|\mathbf{x}_i)) + c_i \log(S_u(t_i|\mathbf{x}_i))\right] \\ &+ \sum_{i=1}^{n} \left[d_i \log p_1 + (1-d_i)\log(1-p_1)\right] c_i(1-\delta_i)\eta_i \\ &+ \sum_{i=1}^{n} \left[d_i \log p_0 + (1-d_i)\log(1-p_0)\right] (1-c_i)(1-\delta_i)\eta_i, \end{aligned} \tag{3.4}$$

where $h_u(.) = f_u(.)/S_u(.)$ is the hazard function of the failure time distribution of uncured patients.

Let $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ be the parameters related to $\mathbf{x}$ and $\mathbf{z}$ respectively, and $\boldsymbol{\theta}' = \left(\boldsymbol{\beta}', \boldsymbol{\gamma}', p_0, p_1\right)$. The log likelihood function can be separated into three parts: the 1^{st} part contains only the "incidence" parameter $\boldsymbol{\gamma}$ related to covariate $\mathbf{z}$; the 2^{nd} part contains only the "latency" parameter $\boldsymbol{\beta}$ related to covariate $\mathbf{x}$, the third part contains only the sensitivity parameter p_0 and specificity parameter $1-p_1$. The three parts can be maximized separately given $\mathbf{c}$. The EM algorithm can be carried out by the following steps.

Initial value: The EM algorithm starts with an initial value $\boldsymbol{\theta}^{(0)}$.

E-step: E-step in the $(r+1)^{th}$ iteration calculates the expectation of the complete log-likelihood function $l(\theta)$, conditional on the observed data and $\boldsymbol{\theta}^{(r)}$ the estimate of $\boldsymbol{\theta}$ at the r^{th} iteration. This is equivalent to calculating the following conditional expectation,

$$g_i^{(r)} = E\left(c_i|\boldsymbol{\theta}^{(r)}, \mathbf{O}\right) = P(c_i = 1|\boldsymbol{\theta}^{(r)}, \mathbf{O}),$$

which is the r^{th} estimator of the probability of the i^{th} patient being uncured. Since

$$\begin{aligned} P(c_i = 1|d_i = 1, \delta_i = 0, \boldsymbol{\theta}^{(r)}, \mathbf{O}) =& \eta_i \frac{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i) p_1^{(r)}}{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i) p_1^{(r)} + (1-\pi^{(r)}(\mathbf{z}_i)) p_0^{(r)}} \\ &+ (1-\eta_i) \frac{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i)}{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i) + (1-\pi^{(r)}(\mathbf{z}_i))}, \end{aligned}$$

and

$$\begin{aligned} P(c_i = 1|d_i = 0, \delta_i = 0, \boldsymbol{\theta}^{(r)}, \mathbf{O}) =& \eta_i \frac{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i)(1-p_1^{(r)})}{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i)(1-p_1^{(r)}) + (1-\pi^{(r)}(\mathbf{z}_i))(1-p_0^{(r)})} \\ &+ (1-\eta_i) \frac{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i)}{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i) + (1-\pi^{(r)}(\mathbf{z}_i))}, \end{aligned}$$

we have

$$
\begin{aligned}
g_i^{(r)} &= P(c_i = 1|\theta^{(r)}, \mathbf{O}) \\
&= \delta_i + (1-\delta_i) d_i \eta_i \frac{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i) p_1^{(r)}}{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i) p_1^{(r)} + (1-\pi^{(r)}(\mathbf{z}_i)) p_0^{(r)}} \\
&\quad + (1-\delta_i)(1-d_i)\eta_i \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|\mathbf{x}_i)(1-p_1^{(r)})}{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i)(1-p_1^{(r)}) + (1-\pi^{(r)}(\mathbf{z}_i))(1-p_0^{(r)})} \\
&\quad + (1-\delta_i)(1-\eta_i) \frac{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i)}{\pi^{(r)}(\mathbf{z}_i) S_u^{(r)}(t_i|\mathbf{x}_i) + (1-\pi^{(r)}(\mathbf{z}_i))}. \qquad (3.5)
\end{aligned}
$$

M-step: M-step in the $(r+1)^{th}$ iteration maximizes the expected complete log-likelihood function with respect to θ to obtain $\theta^{(r+1)}$, where the expected complete log-likelihood function is the sum of the following three functions:

$$
l_1(\boldsymbol{\gamma}|g_i^{(r)}, \mathbf{O}) = \sum_{i=1}^{n} \left[g_i^{(r)} \log(\pi(\mathbf{z}_i)) + (1-g_i^{(r)}) \log(1-\pi(\mathbf{z}_i)) \right], \qquad (3.6)
$$

$$
l_2(\boldsymbol{\beta}|g_i^{(r)}, \mathbf{O}) = \sum_{i=1}^{n} \left[\delta_i \log(h_u(t_i|\mathbf{x}_i)) + g_i^{(r)} \log(S_u(t_i|\mathbf{x}_i)) \right], \qquad (3.7)
$$

$$
\begin{aligned}
l_3(p_0, p_1|g_i^{(r)}, \mathbf{O}) &= \sum_{i=1}^{n} \left[d_i \log(p_1) + (1-d_i) \log(1-p_1) \right] g_i^{(r)} (1-\delta_i)\eta_i \\
&\quad + \sum_{i=1}^{n} \left[d_i \log(p_0) + (1-d_i) \log(1-p_0) \right] (1-g_i^{(r)})(1-\delta_i)\eta_i. \qquad (3.8)
\end{aligned}
$$

When logistic regression model is used to model $\pi(z_i)$, equation (3.6) is the log-likelihood function of a logistic regression model for values arising from a binomial distribution with the response probability $\pi_i = \exp(\boldsymbol{\gamma}'\mathbf{z}_i)/(1+\exp(\boldsymbol{\gamma}'\mathbf{z}_i))$. It can be maximized by the usual optimization methods such as Newton-Raphson method. This maximization procedure can be carried out in most standard logistic regression program to obtain the estimate of $\boldsymbol{\gamma}$. Other link functions can be similarly handled.

For equation (3.7), if a parametric distribution is used for $f_u(t|\mathbf{x})$, the maximization steps can be carried out using Newton-Raphson method. If a semi-parametric proportional hazards model is used, the maximization can be handled using the methods by Peng (2003) as reviewed in Section 2.3.3. We can first obtain the estimate of $\boldsymbol{\beta}$ by a program for the Cox proportional hazards model that accepts covariates with fixed coefficients (Peng, 2003a). After getting $\boldsymbol{\beta}^{(r+1)}$, $S_0^{(r+1)}$ can be obtained by (2.11).

If a semi-parametric accelerated failure time model is used, maximization of $l_2(\boldsymbol{\beta}|g_i^{(r)}, \mathbf{O})$ involves the joint estimation of $\boldsymbol{\beta}$ and S_0. This maximization may be handled using the grid search of the estimating equations (2.20) as proposed by Li and Taylor (2002). Alternatively, Zhang and Peng (2007) proposed the use of linear programming approach to minimize (2.26) in order to obtain the estimate of $\boldsymbol{\beta}$. After $\boldsymbol{\beta}^{(r+1)}$ is obtained, $S_0^{(r+1)}$ can be estimated by (2.27). Detail of these two estimation methods are reviewed in Section 2.4.

If a non-parametric model is used, the survival function can be estimated by (2.28), following the same approach by Taylor (1995) as reviewed in Section 2.5.

Lastly, for equation (3.8), $p_0^{(r+1)}$ and $p_1^{(r+1)}$ can be obtained explicitly from $l_3(p_0, p_1|g_i^{(r)})$ by using the following updating formula:

$$p_0^{(r+1)} = \frac{\sum_{i=1}^n \left[d_i\eta_i(1-g_i^{(r)})(1-\delta_i)\right]}{\sum_{i=1}^n \left[\eta_i(1-g_i^{(r)})(1-\delta_i)\right]} = \frac{\sum_{i:\delta_i=0 \ \& \ \eta_i=1} \left[d_i(1-g_i^{(r)})\right]}{\sum_{i:\delta_i=0 \ \& \ \eta_i=1}(1-g_i^{(r)})},$$

$$p_1^{(r+1)} = \frac{\sum_{i=1}^n \left[d_i\eta_i g_i^{(r)}(1-\delta_i)\right]}{\sum_{i=1}^n \left[\eta_i g_i^{(r)}(1-\delta_i)\right]} = \frac{\sum_{i:\delta_i=0 \ \& \ \eta_i=1} \left[d_i g_i^{(r)}\right]}{\sum_{i:\delta_i=0 \ \& \ \eta_i=1} g_i^{(r)}}.$$

If the sensitivity and specificity are known from the diagnostic procedure externally, the estimations of p_0 and p_1 are not needed.

Iteration: The algorithm is iterated until $\left\|\theta^{(r+1)} - \theta^{(r)}\right\|$ is sufficiently small.

The EM algorithm does not provide the variance estimates of the estimated parameters directly since the information matrix is not available in the EM iteration steps. The approximation method based on Louis method (1982) may be used. As suggested in Peng (2003a), "The accuracy of the approximation is not clear and it cannot be obtained from the iterations of the EM algorithm". We will follow the suggestion in Peng (2003a) to use bootstrap methods for the variance estimates of the estimated parameters.

Chapter 4

Evaluations of the Proposed Models

4.1 Asymptotic Relative Efficiency of the Proposed Estimator

In this section, we assume logit link for the incidence part, exponential distribution for the latency part, and with known p_0 and p_1. Specifically, we assume:

- $\log\left(\frac{\pi(\mathbf{z}_i)}{1-\pi(\mathbf{z}_i)}\right) = \boldsymbol{\gamma}^T\mathbf{z}_i$, where $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, ..., \gamma_k)^T$ is a $1\times(k+1)$ parameter vector, and $\mathbf{z}_i = (z_{i,0}, z_{i,1}, ..., z_{i,k})$ is a $(k+1)\times 1$ covariate vector. $z_{i,0} = 1$.
- $f_u(t_i|\mathbf{x}_i) = h(\mathbf{x}_i)e^{-h(\mathbf{x}_i)t_i}$, and $h(\mathbf{x}_i) = e^{\boldsymbol{\beta}^T\mathbf{x}_i}$, where $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_m)^T$ is a $1\times(m+1)$ parameter vector, and $\mathbf{x}_i = (x_{i,0}, x_{i,1}, ..., x_{i,m})$ is a $(m+1)\times 1$ covariate vector. $x_{i,0} = 1$.
- Known p_0, p_1. For a valid diagnostic procedure, it also assumes that $p_0 \geq p_1$.

For convenience, for all the derivation in this section, we denote

$$\pi_i = \pi(\mathbf{z}_i),$$

and

$$h_i = h(\mathbf{x}_i).$$

The observed likelihood in (3.1) can be written as:

$$\begin{aligned} L_{obs}(\boldsymbol{\theta};\mathbf{O}) &= \prod_{i=1}^{n}\left[\pi_i h_i e^{-h_i t_i}\right]^{\delta_i}\Big[{p_1}^{d_i}(1-p_1)^{1-d_i}\pi_i e^{-h_i t_i} \\ &\qquad + p_0^{d_i}(1-p_0)^{1-d_i}(1-\pi_i)\Big]^{1-\delta_i}. \quad (4.1) \end{aligned}$$

The log-likelihood is:

$$\begin{aligned} l &= \log(L_{obs}(\boldsymbol{\theta}; \mathbf{O})) \\ &= \sum_{i=1}^{n} \delta_i \left[\log(\pi_i) + \log(h_i) - h_i t_i\right] \\ &\quad + \sum_{i=1}^{n} (1-\delta_i) \log\left[{p_1}^{d_i}(1-p_1)^{1-d_i}\pi_i e^{-h_i t_i} + p_0^{d_i}(1-p_0)^{1-d_i}(1-\pi_i)\right]. \end{aligned} \tag{4.2}$$

The score functions are:

$$\begin{aligned} \frac{\partial l}{\partial \boldsymbol{\gamma}} &= \frac{\partial l}{\partial \pi_i}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \\ &= \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \left[\frac{\delta_i}{\pi_i} + (1-\delta_i)\frac{{p_1}^{d_i}(1-p_1)^{1-d_i}e^{-h_i t_i} - p_0^{d_i}(1-p_0)^{1-d_i}}{{p_1}^{d_i}(1-p_1)^{1-d_i}\pi_i e^{-h_i t_i} + p_0^{d_i}(1-p_0)^{1-d_i}(1-\pi_i)}\right], \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} \frac{\partial l}{\partial \boldsymbol{\beta}} &= \frac{\partial l}{\partial h_i}\frac{\partial h_i}{\partial \boldsymbol{\beta}} \\ &= \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \left[\frac{\delta_i}{h_i} - \delta_i t_i \right. \\ &\qquad \left. - (1-\delta_i)\frac{{p_1}^{d_i}(1-p_1)^{1-d_i}\pi_i t_i e^{-h_i t_i}}{{p_1}^{d_i}(1-p_1)^{1-d_i}\pi_i e^{-h_i t_i} + p_0^{d_i}(1-p_0)^{1-d_i}(1-\pi_i)}\right], \end{aligned} \tag{4.4}$$

Define

$$\begin{aligned} a_i &= p_0^{d_i}(1-p_0)^{1-d_i}, \\ b_i &= {p_1}^{d_i}(1-p_1)^{1-d_i}, \end{aligned}$$

and

$$v_i = \frac{b_i}{a_i}.$$

(4.3) and (4.4) can be simplified to

$$\frac{\partial l}{\partial \boldsymbol{\gamma}} = \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \left[\frac{\delta_i}{\pi_i} + (1-\delta_i)\frac{v_i e^{-h_i t_i} - 1}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i}\right], \tag{4.5}$$

$$\frac{\partial l}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \left[\frac{\delta_i}{h_i} - \delta_i t_i - (1-\delta_i)\frac{v_i \pi_i t_i e^{-h_i t_i}}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i}\right]. \tag{4.6}$$

The entries of the observed information matrix are

$$\begin{aligned}
\mathbf{I}_{11} &= -\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} \\
&= \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} \left[\frac{\delta_i}{\pi_i^2} + (1-\delta_i) \frac{\left(v_i e^{-h_i t_i} - 1\right)^2}{\left(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right] \\
&\quad - \sum_{i=1}^{n} \frac{\partial^2 \pi_i}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} \left[\frac{\delta_i}{\pi_i} + (1-\delta_i) \frac{v_i e^{-h_i t_i} - 1}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i} \right], \qquad (4.7) \\
\mathbf{I}_{22} &= -\frac{\partial^2 l}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \\
&= \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \left[\frac{\delta_i}{h_i^2} - (1-\delta_i) \frac{v_i \pi_i \left(1-\pi_i\right) t_i^2 e^{-h_i t_i}}{\left(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right] \\
&\quad - \sum_{i=1}^{n} \frac{\partial^2 h_i}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \left[\frac{\delta_i}{h_i} - \delta_i t_i - (1-\delta_i) \frac{v_i \pi_i t_i e^{-h_i t_i}}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i} \right], \qquad (4.8) \\
\mathbf{I}_{12} &= -\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\beta}^T} \\
&= \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \left[(1-\delta_i) \frac{v_i t_i e^{-h_i t_i}}{\left(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right]. \qquad (4.9)
\end{aligned}$$

For any γ_m and γ_n, and for observation i, since

$$\pi_i = \frac{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}}{1 + e^{\boldsymbol{\gamma}^T \mathbf{z}_i}},$$

the first order partial derivatives are

$$\begin{aligned}
\frac{\partial \pi_i}{\partial \gamma_m} &= \frac{z_{i,m} e^{\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1 + e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^2}, \\
\frac{\partial \pi_i}{\partial \gamma_m} \frac{\partial \pi_i}{\partial \gamma_n} &= \frac{z_{i,m} z_{i,n} e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1 + e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^4}, \\
\left(\frac{\partial \pi_i}{\partial \gamma_m}\right)^2 &= \frac{z_{i,m}^2 e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1 + e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^4}.
\end{aligned}$$

The second order partial derivatives are

$$\begin{aligned}\frac{\partial^2 \pi_i}{\partial \gamma_m^2} &= \frac{-2z_{i,m}^2 e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^3} + \frac{z_{i,m}^2 e^{\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^2} \\ &= \frac{z_{i,m}^2 e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\left(1-e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^3} \\ &= \frac{z_{i,m}^2 e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\left(1-e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}\right)}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^4} \\ &= \frac{1-e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}} \left(\frac{\partial \pi_i}{\partial \gamma_m}\right)^2, \end{aligned} \tag{4.10}$$

and

$$\begin{aligned}\frac{\partial^2 \pi_i}{\partial \gamma_m \partial \gamma_n} &= \frac{-2z_{i,m} z_{i,n} e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^3} + \frac{z_{i,m} z_{i,n} e^{\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^2} \\ &= \frac{z_{i,m} z_{i,n} e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\left(1-e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^3} \\ &= \frac{z_{i,m} z_{i,n} e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\left(1-e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}\right)}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^4} \\ &= \frac{1-e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}} \frac{\partial \pi_i}{\partial \gamma_m} \frac{\partial \pi_i}{\partial \gamma_n}. \end{aligned} \tag{4.11}$$

From (4.10) and (4.11), we have

$$\frac{\partial^2 \pi_i}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} = \frac{1-e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T}. \tag{4.12}$$

Similarly for any β_m and β_n, and for covariate $\mathbf{x}_i$, the first order partial derivatives are

$$\begin{aligned}\frac{\partial h_i}{\partial \beta_m} &= x_{i,m} e^{\boldsymbol{\beta}^T \mathbf{x}_i}, \\ \frac{\partial h_i}{\partial \beta_m} \frac{\partial h_i}{\partial \beta_n} &= x_{i,m} x_{i,n} e^{2\boldsymbol{\beta}^T \mathbf{x}_i}, \\ \left(\frac{\partial h_i}{\partial \beta_m}\right)^2 &= x_{i,m}^2 e^{2\boldsymbol{\beta}^T \mathbf{x}_i}. \end{aligned}$$

The second order partial derivatives are

$$\frac{\partial^2 h_i}{\partial \beta_m^2} = x_{i,m}^2 e^{\boldsymbol{\beta}^T \mathbf{x}_i} = \frac{1}{h_i}\left(\frac{\partial h_i}{\partial \beta_m}\right)^2, \tag{4.13}$$

$$\frac{\partial^2 h_i}{\partial \beta_m \partial \beta_n} = x_{i,m} x_{i,n} e^{\boldsymbol{\beta}^T \mathbf{x}_i} = \frac{1}{h_i} \frac{\partial h_i}{\partial \beta_m} \frac{\partial h_i}{\partial \beta_n}. \tag{4.14}$$

From (4.13) and (4.14), we have

$$\frac{\partial^2 h_i}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} = \frac{1}{h_i} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T}. \tag{4.15}$$

We will then further re-write (4.7), and (4.8) to

$$\begin{aligned} \mathbf{I}_{11} = & \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} \left[\frac{\delta_i}{\pi_i^2} + (1-\delta_i) \frac{\left(v_i e^{-h_i t_i} - 1\right)^2}{\left(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right] \\ & - \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} \frac{1 - e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}} \left[\frac{\delta_i}{\pi_i} + (1-\delta_i) \frac{v_i e^{-h_i t_i} - 1}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i} \right], \end{aligned} \tag{4.16}$$

and

$$\begin{aligned} \mathbf{I}_{22} = & \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \left[\frac{\delta_i}{h_i^2} - (1-\delta_i) \frac{v_i \pi_i \left(1-\pi_i\right) t_i^2 e^{-h_i t_i}}{\left(v_i \pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right] \\ & - \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \frac{1}{h_i} \left[\frac{\delta_i}{h_i} - \delta_i t_i - (1-\delta_i) \frac{v_i \pi_i t_i e^{-h_i t_i}}{v_i \pi_i e^{-h_i t_i} + 1 - \pi_i} \right]. \end{aligned} \tag{4.17}$$

Similarly, if no diagnostic information is used, we have (need only to set $v_i \equiv 1$, or $p_0 = p_1 = 0.5$ in (4.16), (4.17), and (4.9))

$$\begin{aligned} \mathbf{J}_{11} = & -\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} \\ = & \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} \left[\frac{\delta_i}{\pi_i^2} + (1-\delta_i) \frac{\left(e^{-h_i t_i} - 1\right)^2}{\left(\pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right] \\ & - \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} \frac{1 - e^{2\boldsymbol{\gamma}^T \mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}} \left[\frac{\delta_i}{\pi_i} + (1-\delta_i) \frac{e^{-h_i t_i} - 1}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right], \end{aligned} \tag{4.18}$$

$$\begin{aligned} \mathbf{J}_{22} = & -\frac{\partial^2 l}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \\ = & \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \left[\frac{\delta_i}{h_i^2} - (1-\delta_i) \frac{\pi_i \left(1-\pi_i\right) t_i^2 e^{-h_i t_i}}{\left(\pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right] \\ & - \sum_{i=1}^{n} \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \frac{1}{h_i} \left[\frac{\delta_i}{h_i} - \delta_i t_i - (1-\delta_i) \frac{\pi_i t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right], \end{aligned} \tag{4.19}$$

$$\begin{aligned} \mathbf{J}_{12} = & -\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\beta}^T} \\ = & \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \left[(1-\delta_i) \frac{t_i e^{-h_i t_i}}{\left(\pi_i e^{-h_i t_i} + 1 - \pi_i\right)^2} \right]. \end{aligned} \tag{4.20}$$

Denote $\mathbf{T} = \{t_i, i = 1, 2, ..., n\}$, $\mathbf{V} = \{(\delta_i, d_i), i = 1, 2, ..., n\}$. To obtain the information matrix, we will take expectation of $\mathbf{I}_{ij}$ and $\mathbf{J}_{ij}$ with respect to $\mathbf{O} = \{\mathbf{T}, \mathbf{V}\}$. We have the following results.

Lemma 1. *Denote* $\mathbf{I}_{12}^{(i)}$ *and* $\mathbf{J}_{12}^{(i)}$ *as the* i^{th} *summand of* $\mathbf{I}_{12}$ *and* $\mathbf{J}_{12}$*, respectively. Then*

$$\begin{aligned}\Delta_{12}^{(i)} &= E_{\mathbf{O}}\left(\mathbf{I}_{12}^{(i)}\right) - E_{\mathbf{O}}\left(\mathbf{J}_{12}^{(i)}\right) \\ &= -\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}}\left\{\frac{t_i \pi_i e^{-2h_i t_i}\left(1-\pi_i\right)}{\pi_i e^{-ht_i}+\left(1-\pi_i\right)}\varphi_i(p_0,p_1)\right\}\end{aligned}$$

where

$$\varphi_i(p_0,p_1) = \frac{\left(p_0-p_1\right)^2}{\left[(1-p_1)\pi_i e^{-h_i t_i}+\left(1-p_0\right)\left(1-\pi_i\right)\right]\left[p_1\pi_i e^{-h_i t_i}+p_0\left(1-\pi_i\right)\right]}. \tag{4.21}$$

Proof.

Since

$$\begin{aligned}E_{\mathbf{O}}\left(\mathbf{I}_{12}^{(i)}\right) &= E_{\mathbf{O}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\left(1-\delta_i\right)\frac{v_i t_i e^{-h_i t_i}}{\left(v_i\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\} \\ &= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\left(1-\delta_i\right)\frac{v_i t_i e^{-h_i t_i}}{\left(v_i\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\}\right\} \\ &= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\frac{\left(1-\delta_i\right)\left(1-d_i\right)(1-p_1)(1-p_0)t_i e^{-h_i t_i}}{\left[(1-p_1)\pi_i e^{-h_i t_i}+\left(1-p_0\right)\left(1-\pi_i\right)\right]^2}\right]\right\}\right\} \\ &\quad + E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\frac{\left(1-\delta_i\right)d_i p_1 p_0 t_i e^{-h_i t_i}}{\left[p_1\pi_i e^{-h_i t_i}+p_0\left(1-\pi_i\right)\right]^2}\right]\right\}\right\} \\ &= E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}P(\delta_i=0,\, d_i=0|t_i)\left[\frac{(1-p_1)(1-p_0)t_i e^{-h_i t_i}}{\left[(1-p_1)\pi_i e^{-h_i t_i}+\left(1-p_0\right)\left(1-\pi_i\right)\right]^2}\right]\right\} \\ &\quad + E_{\mathbf{T}}\left\{\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}P(\delta_i=0,\, d_i=1|t_i)\left[\frac{p_1 p_0 t_i e^{-h_i t_i}}{\left[p_1\pi_i e^{-h_i t_i}+p_0\left(1-\pi_i\right)\right]^2}\right]\right\},\end{aligned} \tag{4.22}$$

for each i, we have

$$P(\delta_i=0,\, d_i=0|t_i) = (1-p_1)\pi_i e^{-h_i t_i}+\left(1-p_0\right)\left(1-\pi_i\right), \tag{4.23}$$

and

$$P(\delta_i=0,\, d_i=1|t_i) = p_1\pi_i e^{-h_i t_i}+p_0\left(1-\pi_i\right). \tag{4.24}$$

Plug (4.23) and (4.24) into (4.22), we get

$$
\begin{aligned}
E_{\mathbf{O}}\left(\mathbf{I}_{12}^{(i)}\right) &= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ \left[(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i) \right] \right. \\
&\qquad \left. \times \left[\frac{(1-p_1)(1-p_0) t_i e^{-h_i t_i}}{\left[(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i) \right]^2} \right] \right\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ \left[p_1 \pi_i e^{-h_i t_i} + p_0 (1-\pi_i) \right] \right. \\
&\qquad \left. \times \left[\frac{p_1 p_0 t_i e^{-h_i t_i}}{\left[p_1 \pi_i e^{-h_i t_i} + p_0 (1-\pi_i) \right]^2} \right] \right\} \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left[\frac{(1-p_1)(1-p_0) t_i e^{-h_i t_i}}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)} \right] \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left[\frac{p_1 p_0 t_i e^{-h_i t_i}}{p_1 \pi_i e^{-h_i t_i} + p_0 (1-\pi_i)} \right] \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ t_i e^{-h_i t_i} \left[\frac{(1-p_1)(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i} + (1-p_0)(1-\pi_i)} \right. \right. \\
&\qquad \left. \left. + \frac{p_1 p_0}{p_1 \pi_i e^{-h_i t_i} + p_0 (1-\pi_i)} \right] \right\}. \qquad (4.25)
\end{aligned}
$$

Similarly, we can get

$$
E_{\mathbf{O}}\left(\mathbf{J}_{12}^{(i)}\right) = \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ \left[\frac{t_i e^{-h_i t_i}}{\pi_i e^{-h t_i} + (1-\pi_i)} \right] \right\}. \qquad (4.26)
$$

From (4.25) and (4.26), we have

$$\begin{aligned}
\Delta_{12}^{(i)} &= E_{\mathbf{O}}\left(\mathbf{I}_{12}^{(i)}\right) - E_{\mathbf{O}}\left(\mathbf{J}_{12}^{(i)}\right) \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ t_i e^{-ht_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 \left.\left. + \frac{p_1 p_0}{p_1 \pi_i e^{-h_i t_i} + p_0 (1-\pi_i)} \right] \right\} \\
&\quad - \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left[\frac{t_i e^{-h_i t_i}}{\pi_i e^{-ht_i} + (1-\pi_i)} \right] \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} 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 \left.\left. - \frac{1-p_0}{\pi_i e^{-ht_i} + (1-\pi_i)} \right] \right\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} 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)} \right.\right. \\
&\qquad \left.\left. - \frac{p_0}{\pi_i e^{-ht_i} + (1-\pi_i)} \right] \right\} \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ \frac{t_i e^{-h_i t_i}}{\pi_i e^{-ht_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}^T} E_{\mathbf{T}} \left\{ \frac{t_i e^{-h_i t_i}}{\pi_i e^{-ht_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}^T} E_{\mathbf{T}} \left\{ \frac{t_i \pi_i e^{-2h_i t_i} (1-\pi_i)}{\pi_i e^{-ht_i} + (1-\pi_i)} \right. \\
&\qquad \left. \times \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\}.
\end{aligned}$$

Let

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

then $\Delta_{12}^{(i)}$ can be rewritten as

$$\Delta_{12}^{(i)} = -\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ \frac{t_i \pi_i e^{-2h_i t_i} (1-\pi_i)}{\pi_i e^{-ht_i} + (1-\pi_i)} \varphi_i(p_0, p_1) \right\}.$$

□

Lemma 2. *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*

$$\begin{aligned}
\Delta_{11}^{(i)} &= E_{\mathbf{O}}\left(\mathbf{I}_{11}^{(i)}\right) - E_{\mathbf{O}}\left(\mathbf{J}_{11}^{(i)}\right) \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \frac{\pi_i^2 e^{-2h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \varphi_i(p_0, p_1) \right\}
\end{aligned}$$

where

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

Notice that $\varphi_i(p_0,p_1)$ is in the same as that in (4.21).

Proof.

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

the third term in this expression is,

$$\begin{aligned}
&E_{\mathbf{O}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left[(1-\delta_i)\,\frac{v_i e^{-h_i t_i}-1}{v_i\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\}\\
&= E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\right.\right.\\
&\qquad\qquad\left.\left.\times\left[(1-\delta_i)\left(1-d_i\right)\frac{(1-p_1)e^{-h_i t_i}-(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i}+(1-p_0)\left(1-\pi_i\right)}\right]\right\}\right\}\\
&\quad + E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{(1-\delta_i)\,d_i\frac{p_1 e^{-h_i t_i}-p_0}{p_1\pi_i e^{-h_i t_i}+p_0\left(1-\pi_i\right)}\right\}\right\}\\
&= E_{\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\right.\\
&\qquad\qquad\left.\times\left[P(\delta_i=0,\,d_i=0|t_i)\frac{(1-p_1)e^{-h_i t_i}-(1-p_0)}{(1-p_1)\pi_i e^{-h_i t_i}+(1-p_0)\left(1-\pi_i\right)}\right]\right\}\\
&\quad + E_{\mathbf{T}}\left\{P(\delta_i=0,\,d_i=1|t_i)\frac{p_1 e^{-h_i t_i}-p_0}{p_1\pi_i e^{-h_i t_i}+p_0\left(1-\pi_i\right)}\right\}.
\end{aligned}$$

From (4.23) and (4.24),

$$\begin{aligned}
& E_{\mathbf{O}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left[(1-\delta_i)\,\frac{v_i e^{-h_i t_i}-1}{v_i\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\} \\
= \; & E_{\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left[(1-p_1)e^{-h_i t_i}-(1-p_0)\right]\right\} \\
& +E_{\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left[p_1 e^{-h_i t_i}-p_0\right]\right\} \\
= \; & E_{\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left(e^{-h_i t_i}-1\right)\right\}.
\end{aligned}$$

Since this expression does not depend on v_i, we have

$$\begin{aligned}
& E_{\mathbf{O}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left[(1-\delta_i)\,\frac{v_i e^{-h_i t_i}-1}{v_i\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\} \\
= \; & E_{\mathbf{O}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left[(1-\delta_i)\,\frac{e^{-h_i t_i}-1}{\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\Delta_{11}^{(i)} = \; & E_{\mathbf{O}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\left[(1-\delta_i)\,\frac{\left(v_i e^{-h_i t_i}-1\right)^2}{\left(v_i\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\} \\
& -E_{\mathbf{O}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\left[(1-\delta_i)\,\frac{\left(e^{-h_i t_i}-1\right)^2}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^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}^T}\left[\frac{(1-\delta_i)\,(1-d_i)\left[(1-p_1)e^{-h_i t_i}-(1-p_0)\right]^2}{\left[(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)\,(1-p_0)\right]^2}\right]\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}^T}\left[\frac{(1-\delta_i)\,d_i\left[p_1 e^{-h_i t_i}-p_0\right]^2}{\left[p_1\pi_i e^{-h_i t_i}+p_0\,(1-\pi_i)\right]^2}\right]\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}^T}\left[\frac{(1-\delta_i)\left(e^{-h_i t_i}-1\right)^2}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\}\right\} \\
= \; & E_{\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\left[\frac{P(\delta_i=0,\,d_i=0|t_i)\left[(1-p_1)e^{-h_i t_i}-(1-p_0)\right]^2}{\left[(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)\,(1-p_0)\right]^2}\right]\right\} \\
& +E_{\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\left[\frac{P(\delta_i=0,\,d_i=1|t_i)\left[p_1 e^{-h_i t_i}-p_0\right]^2}{\left[p_1\pi_i e^{-h_i t_i}+p_0\,(1-\pi_i)\right]^2}\right]\right\} \\
& -E_{\mathbf{T}}\left\{\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\left[\frac{P(\delta_i=0|t_i)\left(e^{-h_i t_i}-1\right)^2}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\}.
\end{aligned}$$

Again because of (4.23) and (4.24),

$$\begin{aligned}
\Delta_{11}^{(i)} &= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \frac{\left[(1-p_1)e^{-h_i t_i} - (1-p_0)\right]^2}{(1-p_1)\pi_i e^{-h_i t_i} + (1-\pi_i)(1-p_0)} \right\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \frac{\left[p_1 e^{-h_i t_i} - p_0\right]^2}{p_1 \pi_i e^{-h_i t_i} + p_0 (1-\pi_i)} \right\} \\
&\quad - \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \frac{\left(e^{-h_i t_i} - 1\right)^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}^T} E_{\mathbf{T}} \left\{ \left[\frac{\left[(1-p_1)e^{-h_i t_i} - (1-p_0)\right]^2}{(1-p_1)\pi_i e^{-h_i t_i} + (1-\pi_i)(1-p_0)} \right] \right. \\
&\qquad \left. - \left[\frac{(1-p_1)\left(e^{-h_i t_i} - 1\right)^2}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right] \right\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \left[\frac{\left[p_1 e^{-h_i t_i} - p_0\right]^2}{p_1 \pi_i e^{-h_i t_i} + p_0 (1-\pi_i)} \right] \right. \\
&\qquad \left. - \left[\frac{p_1 \left(e^{-h_i t_i} - 1\right)^2}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right] \right\} \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \frac{p_1 - p_0}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right. \\
&\qquad \left. \times \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\} \\
&\quad + \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \left[\frac{p_0 - p_1}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \right. \right. \\
&\qquad \left. \left. \times \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] \right\} \\
&= \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} E_{\mathbf{T}} \left\{ \frac{\pi_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}$$

□

Lemma 3. *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*

$$\begin{aligned}
\Delta_{22}^{(i)} &= E_{\mathbf{O}}\left(\mathbf{I}_{22}^{(i)}\right) - E_{\mathbf{O}}\left(\mathbf{J}_{22}^{(i)}\right) \\
&= \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} E_{\mathbf{T}} \left\{ \frac{\pi_i^2 (1-\pi_i)^2 t_i^2 e^{-2h_i t_i}}{\pi_i e^{-h_i t_i} + 1 - \pi_i} \varphi_i(p_0, p_1) \right\}
\end{aligned}$$

where

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

Notice that $\varphi_i(p_0, p_1)$ *is in the same as that in (4.21).*

Proof.

$$
\begin{aligned}
\Delta_{22}^{(i)} &= E_{\mathbf{O}}\left(\mathbf{I}_{22}^{(i)}\right) - E_{\mathbf{O}}\left(\mathbf{J}_{22}^{(i)}\right) \\
&= -E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[(1-\delta_i)\,\frac{v_i\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{(v_i\pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\} \\
&\quad +E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[(1-\delta_i)\,\frac{\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{(\pi_i e^{-h_i t_i}+1-\pi_i)^2}\right]\right\} \\
&\quad +E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[(1-\delta_i)\,\frac{v_i\pi_i t_i e^{-h_i t_i}}{v_i\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\} \\
&\quad -E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[(1-\delta_i)\,\frac{\pi_i t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\},
\end{aligned}
$$

the third term is,

$$
\begin{aligned}
& E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[(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]\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}^T}\frac{1}{h_i}\left[\frac{(1-\delta_i)\,(1-d_i)\,(1-p_1)\,\pi_i t_i e^{-h_i t_i}}{(1-p_1)\,\pi_i e^{-h_i t_i}+(1-p_0)\,(1-\pi_i)}\right]\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}^T}\frac{1}{h_i}\left[\frac{(1-\delta_i)\,d_i p_1\pi_i t_i e^{-h_i t_i}}{p_1\pi_i e^{-h_i t_i}+p_0\,(1-\pi_i)}\right]\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}^T}\frac{1}{h_i}\left[\frac{(1-\delta_i)\,(1-d_i)\,(1-p_1)\,\pi_i t_i e^{-h_i t_i}}{(1-p_1)\,\pi_i e^{-h_i t_i}+(1-p_0)\,(1-\pi_i)}\right]\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}^T}\frac{1}{h_i}\left[\frac{(1-\delta_i)\,d_i p_1\pi_i t_i e^{-h_i t_i}}{p_1\pi_i e^{-h_i t_i}+p_0\,(1-\pi_i)}\right]\right\}\right\} \\
= \; & E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[\frac{P(\delta_i=0,\,d_i=0|t_i)\,(1-p_1)\,\pi_i t_i e^{-h_i t_i}}{(1-p_1)\,\pi_i e^{-h_i t_i}+(1-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}^T}\frac{1}{h_i}\left[\frac{P(\delta_i=0,\,d_i=1|t_i)p_1\pi_i t_i e^{-h_i t_i}}{p_1\pi_i e^{-h_i t_i}+p_0\,(1-\pi_i)}\right]\right\}.
\end{aligned}
$$

From (4.23) and (4.24),

$$
\begin{aligned}
& E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[(1-\delta_i)\,\frac{v_i\pi_i t_i e^{-h_i t_i}}{v_i\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\} \\
= \; & E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[(1-p_1)\,\pi_i t_i e^{-h_i t_i}\right]\right\}\right\} \\
& +E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[p_1\pi_i t_i e^{-h_i t_i}\right]\right\}\right\} \\
= \; & E_{\mathbf{T}}\left\{E_{\mathbf{V}|\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[\pi_i t_i e^{-h_i t_i}\right]\right\}\right\}.
\end{aligned}
$$

Since this expression does not depend on v_i, we have

$$\begin{aligned} & E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[(1-\delta_i)\,\frac{v_i\pi_i t_i e^{-h_i t_i}}{v_i\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\} \\ = \; & E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\frac{1}{h_i}\left[(1-\delta_i)\,\frac{\pi_i t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}\right]\right\}. \end{aligned}$$

It follows that

$$\begin{aligned} \Delta_{22}^{(i)} = \; & -E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[(1-\delta_i)\,\frac{v_i\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left(v_i\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\} \\ & +E_{\mathbf{O}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[(1-\delta_i)\,\frac{\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^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}^T}\left[\frac{(1-\delta_i)\,(1-d_i)(1-p_1)(1-p_0)\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left[(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)\,(1-p_0)\right]^2}\right]\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}^T}\left[\frac{(1-\delta_i)\,d_i p_1 p_0\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left[p_1\pi_i e^{-h_i t_i}+p_0\,(1-\pi_i)\right]^2}\right]\right\}\right\} \\ & +E_{\mathbf{T}}\left\{E_{\mathbf{V}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\frac{(1-\delta_i)\,\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\}\right\} \\ = \; & -E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\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_i t_i}}{\left[(1-p_1)\pi_i e^{-h_i t_i}+(1-\pi_i)\,(1-p_0)\right]^2}\right]\right\} \\ & -E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\frac{P(\delta_i=0,\,d_i=1|t_i)p_1 p_0\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left[p_1\pi_i e^{-h_i t_i}+p_0\,(1-\pi_i)\right]^2}\right]\right\} \\ & +E_{\mathbf{T}}\left\{\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\left[\frac{P(\delta_i=0|t_i)\pi_i\,(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right]\right\}. \end{aligned}$$

Based on (4.23) and (4.24)

$$\begin{aligned}
\Delta_{22}^{(i)} &= -\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\frac{(1-p_1)(1-p_0)\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{(1-p_1)\pi_ie^{-h_it_i}+(1-\pi_i)\left(1-p_0\right)}\right\} \\
&\quad -\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\frac{p_1p_0\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{p_1\pi_ie^{-h_it_i}+p_0\left(1-\pi_i\right)}\right\} \\
&\quad +\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\frac{\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{\pi_ie^{-h_it_i}+1-\pi_i}\right\} \\
&= -\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\left[\frac{(1-p_1)(1-p_0)\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{(1-p_1)\pi_ie^{-h_it_i}+(1-\pi_i)\left(1-p_0\right)}\right]\right. \\
&\qquad\qquad \left. -\left[\frac{(1-p_0)\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{\pi_ie^{-h_it_i}+1-\pi_i}\right]\right\} \\
&\quad -\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\left[\frac{p_1p_0\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{p_1\pi_ie^{-h_it_i}+p_0\left(1-\pi_i\right)}\right]\right. \\
&\qquad\qquad \left. -\left[\frac{p_0\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{\pi_ie^{-h_it_i}+1-\pi_i}\right]\right\} \\
&= \frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\frac{(1-p_0)\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{\pi_ie^{-h_it_i}+1-\pi_i}\right. \\
&\qquad\qquad \left.\times\frac{\left(p_1-p_0\right)\left(1-\pi_i\right)}{(1-p_1)\pi_ie^{-h_it_i}+(1-\pi_i)\left(1-p_0\right)}\right\} \\
&\quad +\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\frac{p_0\pi_i\left(1-\pi_i\right)t_i^2e^{-h_it_i}}{\pi_ie^{-h_it_i}+1-\pi_i}\frac{\left(p_0-p_1\right)\left(1-\pi_i\right)}{p_1\pi_ie^{-h_it_i}+p_0\left(1-\pi_i\right)}\right\} \\
&= \frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}E_{\mathbf{T}}\left\{\frac{\pi_i^2\left(1-\pi_i\right)^2t_i^2e^{-2h_it_i}}{\pi_ie^{-h_it_i}+1-\pi_i}\varphi_i(p_0,p_1)\right\}.
\end{aligned}$$

□

Note that the expressions of $\Delta_{12}^{(i)}$, $\Delta_{11}^{(i)}$, and $\Delta_{22}^{(i)}$ all involve $\varphi_i(p_0,p_1)$. For function $\varphi_i(p_0,p_1)$, we have the following lemma.

Lemma 4. *For function*

$$\varphi_i(p_0,p_1)=\frac{\left(p_0-p_1\right)^2}{\left[(1-p_1)\pi_ie^{-h_it_i}+\left(1-p_0\right)\left(1-\pi_i\right)\right]\left[p_1\pi_ie^{-h_it_i}+p_0\left(1-\pi_i\right)\right]},$$

if $1\geq p_0\geq p_1\geq 0$*, 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 we hold p_0 as fixed, we can rewrite $\varphi_i(p_0,p_1)$ as

$$\begin{aligned}
\varphi_i(p_0,p_1) &= \frac{\left(p_0-p_1\right)}{\pi_ie^{-h_it_i}}\left[\frac{p_0}{p_1\pi_ie^{-h_it_i}+p_0\left(1-\pi_i\right)}\right. \\
&\qquad\qquad \left.-\frac{1-p_0}{(1-p_1)\pi_ie^{-h_it_i}+\left(1-p_0\right)\left(1-\pi_i\right)}\right].
\end{aligned}$$

Since $p_0 \geq p_1$, 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, we can rewrite $\varphi_i(p_0, p_1)$ 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 $\varphi_i(p_0, p_1)$.

□

With the differences for each entry of the information matrix computed by Lemma 1, Lemma 2, Lemma 3, and the property of the differences established by Lemma 4. We have the following proposition.

Proposition 1. *If we denote $V_{\boldsymbol{\gamma}}^D$ as the asymptotic variance of $\hat{\boldsymbol{\gamma}}$ when the diagnostic procedure is used, and $V_{\boldsymbol{\gamma}}^N$ as the asymptotic variance of $\hat{\boldsymbol{\gamma}}$ when no diagnostic procedure is used; also denote $V_{\boldsymbol{\beta}}^D$ as the asymptotic variance of $\hat{\boldsymbol{\beta}}$ when the diagnostic procedure is used, and $V_{\boldsymbol{\beta}}^N$ as the asymptotic variance of $\hat{\boldsymbol{\beta}}$ 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 the diagonal entries in $V_{\boldsymbol{\gamma}}^D$ and $V_{\boldsymbol{\beta}}^D$ are less than or equal to the corresponding entries in $V_{\boldsymbol{\gamma}}^N$ and $V_{\boldsymbol{\beta}}^N$. This implies the estimates 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 the estimates of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are more efficient when diagnostic information is included. Furthermore, the 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 $x_{i,1}$ is an indicator of 0/1, the variances of $\hat{\gamma}_0$ and $\hat{\beta}_0$ are smaller when the diagnostic procedure is used. This implies the estimates of γ_0 and β_0 are more efficient when diagnostic*

information is included. Furthermore, the 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 $\gamma_{i,1}$ is an indicator of 0/1, the variances of $\hat{\gamma}_0$ and $\hat{\beta}_0$ are smaller when the diagnostic procedure is used. This implies the estimates of γ_0 and β_0 are more efficient when diagnostic information is included. Furthermore, the variance decreases as the sensitivity or specificity increases.*

Proof.

For notational simplicity, let

$$\begin{aligned}
a_{11}^{(i)} &= \left[\frac{\delta_i}{\pi_i^2} + (1-\delta_i)\frac{\left(e^{-h_i t_i}-1\right)^2}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right] \\
&\quad -\frac{1-e^{2\boldsymbol{\gamma}^T\mathbf{z}_i}}{e^{\boldsymbol{\gamma}^T\mathbf{z}_i}}\left[\frac{\delta_i}{\pi_i} + (1-\delta_i)\frac{e^{-h_i t_i}-1}{\pi_i e^{-h_i t_i}+1-\pi_i}\right], \\
a_{22}^{(i)} &= \left[\frac{\delta_i}{h_i^2} - (1-\delta_i)\frac{\pi_i(1-\pi_i)\,t_i^2 e^{-h_i t_i}}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}\right] \\
&\quad -\frac{1}{h_i}\left[\frac{\delta_i}{h_i} - \delta_i t_i - (1-\delta_i)\frac{\pi_i t_i e^{-h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}\right], \\
a_{12}^{(i)} &= (1-\delta_i)\frac{t_i e^{-h_i t_i}}{\left(\pi_i e^{-h_i t_i}+1-\pi_i\right)^2}.
\end{aligned}$$

and

$$\begin{aligned}
d_{11}^{(i)} &= \frac{\pi_i^2 e^{-2h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}, \\
d_{22}^{(i)} &= \frac{\pi_i^2(1-\pi_i)^2 t_i^2 e^{-2h_i t_i}}{\pi_i e^{-h_i t_i}+1-\pi_i}, \\
d_{12}^{(i)} &= \frac{t_i \pi_i e^{-2h_i t_i}(1-\pi_i)}{\pi_i e^{-ht_i}+(1-\pi_i)}.
\end{aligned}$$

Then

$$\begin{aligned}
E_{\mathbf{O}}\left(\mathbf{J}_{11}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n} a_{11}^{(i)}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T}\right), \\
E_{\mathbf{O}}\left(\mathbf{J}_{22}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n} a_{22}^{(i)}\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\right), \\
E_{\mathbf{O}}\left(\mathbf{J}_{12}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n} a_{12}^{(i)}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\right)
\end{aligned}$$

and by Lemma 1, Lemma 2, and Lemma 3,

$$\Delta_{11} = \sum_{i=1}^{n} \Delta_{11}^{(i)} = E_{\mathbf{T}} \left(\sum_{i=1}^{n} d_{11}^{(i)} \varphi_i(p_0, p_1) \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} \right),$$
$$\Delta_{22} = \sum_{i=1}^{n} \Delta_{22}^{(i)} = E_{\mathbf{T}} \left(\sum_{i=1}^{n} d_{22}^{(i)} \varphi_i(p_0, p_1) \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \right),$$
$$\Delta_{12} = \sum_{i=1}^{n} \Delta_{12}^{(i)} = -E_{\mathbf{T}} \left(\sum_{i=1}^{n} d_{12}^{(i)} \varphi_i(p_0, p_1) \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \right)$$

It is obvious from the expression that $a_{12}^{(i)} \geq 0$, $d_{12}^{(i)} \geq 0$, and $E_{\mathbf{O}} \left[a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1) \right] \geq 0$. Also, because of the positive definite property of 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 $p_0 = 1$, $p_1 = 0$, $\varphi_i(p_0, p_1)$ 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

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

and by (4.26)

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

So,

$$\begin{aligned} E_{\mathbf{O}} \left(\mathbf{I}_{12} \right) &= E_{\mathbf{O}} \left(\sum_{i=1}^{n} \left[a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1) \right] \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \right) \\ &= 0. \end{aligned} \tag{4.27}$$

By (4.27), we get

$$\begin{aligned} V_{\boldsymbol{\gamma}}^{D} &= \left\{ E_{\mathbf{O}} \left(\mathbf{I}_{11} \right) - E_{\mathbf{O}} \left(\mathbf{I}_{12} \right) E_{\mathbf{O}}^{-1} \left(\mathbf{I}_{22} \right) E_{\mathbf{O}}^{T} \left(\mathbf{I}_{12} \right) \right\}^{-1} \\ &= \left\{ E_{\mathbf{O}} \left(\sum_{i=1}^{n} \left[a_{11}^{(i)} + d_{11}^{(i)} \varphi_i(p_0, p_1) \right] \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T} \right) \right\}^{-1}, \end{aligned} \tag{4.28}$$

and

$$\begin{aligned} V_{\boldsymbol{\beta}}^{D} &= \left\{ E_{\mathbf{O}} \left(\mathbf{I}_{22} \right) - E_{\mathbf{O}}^{T} \left(\mathbf{I}_{12} \right) E_{\mathbf{O}}^{-1} \left(\mathbf{I}_{11} \right) E_{\mathbf{O}} \left(\mathbf{I}_{12} \right) \right\}^{-1} \\ &= \left\{ E_{\mathbf{O}} \left(\sum_{i=1}^{n} \left[a_{22}^{(i)} + d_{22}^{(i)} \varphi_i(p_0, p_1) \right] \frac{\partial h_i}{\partial \boldsymbol{\beta}} \frac{\partial h_i}{\partial \boldsymbol{\beta}^T} \right) \right\}^{-1}. \end{aligned} \tag{4.29}$$

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

$$\begin{aligned}
& \mathbf{u}^T\left(\mathbf{A}+c\mathbf{w}\mathbf{w}^T\right)^{-1}\mathbf{u} \\
= \; & \mathbf{u}^T\left(\mathbf{A}^{-1}-\frac{1}{\frac{1}{c}+\mathbf{w}^T\mathbf{A}^{-1}\mathbf{w}}\mathbf{A}^{-1}\mathbf{w}\mathbf{w}^T\mathbf{A}^{-1}\right)\mathbf{u} \\
= \; & \mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}-\frac{1}{\frac{1}{c}+\mathbf{w}^T\mathbf{A}^{-1}\mathbf{w}}\mathbf{u}^T\mathbf{A}^{-1}\mathbf{w}\mathbf{w}^T\mathbf{A}^{-1}\mathbf{u} \\
= \; & \mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}-\frac{1}{\frac{1}{c}+\mathbf{w}^T\mathbf{A}^{-1}\mathbf{w}}\left(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{w}\right)^2 \\
\leq \; & \mathbf{u}^T\mathbf{A}^{-1}\mathbf{u} \qquad (4.30)
\end{aligned}$$

By adding $\Delta_{11}^{(i)}$ one at a time, for any $\mathbf{u}$, we have

$$\mathbf{u}^T V_{\boldsymbol{\gamma}}^D \mathbf{u} \leq \mathbf{u}^T V_{\boldsymbol{\gamma}}^N \mathbf{u}$$

Taking $\mathbf{u}_i$ as

$$u_{ij}=0, \text{ if } j\neq i,$$

and

$$u_{ij}=1, \text{ if } j=i,$$

we can conclude all the diagonal entries of $V_{\boldsymbol{\gamma}}^D$ are less than or equal to the corresponding diagonal entries of $V_{\boldsymbol{\gamma}}^N$. Smaller diagonal entries indicate higher efficiency, hence the estimate of $\boldsymbol{\gamma}$ with diagnostic information included is more efficient than that without.

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

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

$$\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} = C_{\gamma_0},$$

and

$$\frac{\partial h_i}{\partial \boldsymbol{\beta}} = C_{\beta_0}.$$

Then

$$\begin{aligned} E_{\mathbf{O}}\left(\mathbf{I}_{11}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{11}^{(i)}+d_{11}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}^T}\right) \\ &= C_{\gamma_0}^2 E_{\mathbf{O}}\left(a_{11}^{(1)}+d_{11}^{(1)}\varphi_1(p_0,p_1)\right), \end{aligned} \tag{4.31}$$

$$\begin{aligned} E_{\mathbf{O}}\left(\mathbf{I}_{12}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{12}^{(i)}-d_{12}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial h_i}{\partial\boldsymbol{\beta}^T}\right) \\ &= C_{\gamma_0}C_{\beta_0}E_{\mathbf{O}}\left(a_{12}^{(1)}-d_{12}^{(1)}\varphi_1(p_0,p_1)\right), \end{aligned} \tag{4.32}$$

and

$$\begin{aligned} E_{\mathbf{O}}\left(\mathbf{I}_{22}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{22}^{(i)}+d_{22}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial h_i}{\partial\boldsymbol{\beta}}\frac{\partial h_i}{\partial\boldsymbol{\beta}^T}\right) \\ &= C_{\beta_0}^2 E_{\mathbf{O}}\left(a_{22}^{(1)}+d_{22}^{(1)}\varphi_1(p_0,p_1)\right). \end{aligned} \tag{4.33}$$

From (4.31), (4.32), and (4.33),

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

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

Similarly,

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

$\left[V_{\boldsymbol{\beta}}^{D}\right]^{-1}$ is also an increasing function of p_0, and decreasing functions of p_1. Larger $\left[V_{\boldsymbol{\beta}}^{D}\right]^{-1}$ lead to smaller $V_{\boldsymbol{\beta}}^{D}$. Thus the efficiency of the estimate $\hat{\boldsymbol{\beta}}$ increases as either specificity or sensitivity increases, , and the estimate of $\boldsymbol{\beta}$ with diagnostic information included is more efficient than that without diagnostic information included.

Proof of Case 3, since $\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}$ is the same for all subjects because $\boldsymbol{\gamma} = (\gamma_0)$, we can denote it as an unknown constant C_{γ_0}. For $\boldsymbol{\beta} = (\beta_0, \beta_1)$, we have

$$\begin{aligned}
\frac{\partial h_i}{\partial \boldsymbol{\beta}} &= h_i \begin{bmatrix} 1 \\ x_{i,1} \end{bmatrix} \\
&= h_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} I(x_{i,1}=0) + h_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} I(x_{i,1}=1),
\end{aligned}$$

where h_0 and h_1 are the shorthand notation of $h_i(x_{i,1}=0)$ and $h_i(x_{i,1}=1)$.

Let

$$\begin{aligned}
b_{11,0} &= E_{\mathbf{O}}\left(a_{11}^{(i)} + d_{11}^{(i)}\varphi_i(p_0,p_1)|x_{i,1}=0\right), \\
b_{11,1} &= E_{\mathbf{O}}\left(a_{11}^{(i)} + d_{11}^{(i)}\varphi_i(p_0,p_1)|x_{i,1}=1\right), \\
b_{12,0} &= E_{\mathbf{O}}\left(a_{12}^{(i)} - d_{12}^{(i)}\varphi_i(p_0,p_1)|x_{i,1}=0\right), \\
b_{12,1} &= E_{\mathbf{O}}\left(a_{12}^{(i)} - d_{12}^{(i)}\varphi_i(p_0,p_1)|x_{i,1}=1\right), \\
b_{22,0} &= E_{\mathbf{O}}\left(a_{22}^{(i)} + d_{22}^{(i)}\varphi_i(p_0,p_1)|x_{i,1}=0\right), \\
b_{22,1} &= E_{\mathbf{O}}\left(a_{22}^{(i)} + d_{22}^{(i)}\varphi_i(p_0,p_1)|x_{i,1}=1\right).
\end{aligned}$$

Also assume there are n_0 observations with $x_{i,1}=0$, and n_1 observations with $x_{i,1}=1$, by i.i.d. property when the covariates are the same, then

$$\begin{aligned}
E_{\mathbf{O}}\left(\mathbf{I}_{11}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{11}^{(i)} + d_{11}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T}\right) \\
&= C_{\gamma_0}^2\left(n_1 b_{11,1} + n_0 b_{11,0}\right), \qquad (4.34)
\end{aligned}$$

$$\begin{aligned}
E_{\mathbf{O}}\left(\mathbf{I}_{12}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{12}^{(i)} - d_{12}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\right) \\
&= C_{\gamma_0}\left\{n_1 b_{12,1} h_1 \begin{bmatrix} 1 & 1 \end{bmatrix} + n_0 b_{12,0} h_0 \begin{bmatrix} 1 & 0 \end{bmatrix}\right\} \\
&= C_{\gamma_0}\begin{bmatrix} n_1 b_{12,1} h_1 + n_0 b_{12,0} h_0 & n_1 b_{12,1} h_1 \end{bmatrix}, \qquad (4.35)
\end{aligned}$$

and

$$\begin{aligned} E_{\mathbf{O}}\left(\mathbf{I}_{22}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{22}^{(i)}+d_{22}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial h_i}{\partial \boldsymbol{\beta}}\frac{\partial h_i}{\partial \boldsymbol{\beta}^T}\right) \\ &= n_1 b_{22,1} h_1^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + n_0 b_{22,0} h_0^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} n_1 b_{22,1} h_1^2 + n_0 b_{22,0} h_0^2 & n_1 b_{22,1} h_1^2 \\ n_1 b_{22,1} h_1^2 & n_1 b_{22,1} h_1^2 \end{bmatrix}. \end{aligned}$$

As a result,

$$E_{\mathbf{O}}^{-1}\left(\mathbf{I}_{22}\right) = \begin{bmatrix} \frac{1}{n_0 b_{22,0} h_0^2} & -\frac{1}{n_0 b_{22,0} h_0^2} \\ -\frac{1}{n_0 b_{22,0} h_0^2} & \frac{1}{n_0 b_{22,0} h_0^2} + \frac{1}{n_1 b_{22,1} h_1^2} \end{bmatrix}. \tag{4.36}$$

From (4.34), (4.35), and (4.36),

$$\begin{aligned} \left[V_{\boldsymbol{\gamma}}^D\right]^{-1} &= E_{\mathbf{O}}\left(\mathbf{I}_{11}\right) - E_{\mathbf{O}}\left(\mathbf{I}_{12}\right) E_{\mathbf{O}}^{-1}\left(\mathbf{I}_{22}\right) E_{\mathbf{O}}^{T}\left(\mathbf{I}_{12}\right) \\ &= C_{\gamma_0}^2\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - C_{\gamma_0}^2 \begin{bmatrix} n_1 b_{12,1} h_1 + n_0 b_{12,0} h_0, & n_1 b_{12,1} h_1 \end{bmatrix} \\ &\qquad \times \begin{bmatrix} \frac{1}{n_0 b_{22,0} h_0^2} & -\frac{1}{n_0 b_{22,0} h_0^2} \\ -\frac{1}{n_0 b_{22,0} h_0^2} & \frac{1}{n_0 b_{22,0} h_0^2} + \frac{1}{n_1 b_{22,1} h_1^2} \end{bmatrix} \begin{bmatrix} n_1 b_{12,1} h_1 + n_0 b_{12,0} h_0 \\ n_1 b_{12,1} h_1 \end{bmatrix} \\ &= C_{\gamma_0}^2 \left[\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \frac{\left(n_0 b_{12,0} h_0\right)^2}{n_0 b_{22,0} h_0^2} - \frac{\left(n_1 b_{12,1} h_1\right)^2}{n_1 b_{22,1} h_1^2}\right] \\ &= C_{\gamma_0}^2 \left[\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \frac{n_0 b_{12,0}^2}{b_{22,0}} - \frac{n_1 b_{12,1}^2}{b_{22,1}}\right]. \end{aligned}$$

Since $b_{11,1}$, $b_{11,0}$, $b_{22,0}$, $b_{22,1}$ are increasing functions of p_0, and decreasing functions of p_1 through their dependence of $\varphi_i(p_0,p_1)$, and $b_{12,1}$, $b_{12,0}$ are decreasing functions of p_0, and increasing functions of p_1 through their dependence of $\varphi_i(p_0,p_1)$, $\left[V_{\boldsymbol{\gamma}}^D\right]^{-1}$ is an increasing function of p_0, and decreasing functions of p_1. Larger $\left[V_{\boldsymbol{\gamma}}^D\right]^{-1}$ lead to smaller $V_{\boldsymbol{\gamma}}^D$. Thus the efficiency of the estimate $\hat{\boldsymbol{\gamma}}$ increases as either specificity or sensitivity increases, and the estimate of $\boldsymbol{\gamma}$ with diagnostic information included is more efficient than that without diagnostic information included.

For $V_{\boldsymbol{\beta}}^D$, it can be computed as

$$\frac{1}{C_{\gamma_0}^2 G}\begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{bmatrix},$$

where

$$
\begin{aligned}
G &= \frac{1}{\left(n_1 b_{22,1} h_1^2\right)\left(n_0 b_{22,0} h_0^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \left(n_1 b_{22,1} h_1^2\right)\left(n_0 b_{12,0} h_0\right)^2 - \left(n_0 b_{22,0} h_0^2\right)\left(n_1 b_{12,1} h_1\right)^2}, \\
V_{11} &= \left(n_1 b_{22,1} h_1^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \left(n_1 b_{12,1} h_1\right)^2, \\
V_{22} &= \left(n_1 b_{22,1} h_1^2 + n_0 b_{22,0} h_0^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \left(n_1 b_{12,1} h_1 + n_0 b_{12,0} h_0\right)^2, \\
V_{12} &= -\left(n_1 b_{22,1} h_1^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) + \left(n_1 b_{12,1} h_1 + n_0 b_{12,0} h_0\right)\left(n_1 b_{12,1} h_1\right).
\end{aligned}
$$

For the variance of $\hat{\beta}_0$, we have

$$
\begin{aligned}
& C_{\gamma_0}^2 V_{\beta_0}^D \\
= & \frac{\left(n_1 b_{22,1} h_1^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \left(n_1 b_{12,1} h_1\right)^2}{\left(n_1 b_{22,1} h_1^2\right)\left(n_0 b_{22,0} h_0^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \left(n_1 b_{22,1} h_1^2\right)\left(n_0 b_{12,0} h_0\right)^2 - \left(n_0 b_{22,0} h_0^2\right)\left(n_1 b_{12,1} h_1\right)^2} \\
= & \frac{1}{\left(n_0 b_{22,0} h_0^2\right) - \frac{\left(n_1 b_{22,1} h_1^2\right)\left(n_0 b_{12,0} h_0\right)^2}{\left(n_1 b_{22,1} h_1^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \left(n_1 b_{12,1} h_1\right)^2}} \\
= & \frac{1}{\left(n_0 b_{22,0} h_0^2\right) - \frac{\left(n_0 b_{12,0} h_0\right)^2}{\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \frac{\left(n_1 b_{12,1} h_1\right)^2}{\left(n_1 b_{22,1} h_1^2\right)}}}
\end{aligned}
$$

Because $b_{11,1}$, $b_{11,0}$, $b_{22,0}$, $b_{22,1}$ are increasing functions of p_0, and decreasing functions of p_1 through their dependence of $\varphi_i(p_0, p_1)$, and $b_{12,1}$, $b_{12,0}$ are decreasing functions of p_0, and increasing functions of p_1 through their dependence of $\varphi_i(p_0, p_1)$, we can infer that

$$
\begin{aligned}
& \text{Increasing } p_0\text{, and/or decreasing } p_1 \\
\Rightarrow\ & \text{Increasing } b_{11,1}, b_{11,0}, b_{22,0}, b_{22,1}\text{, and decreasing } b_{12,1}, b_{12,0} \\
\Rightarrow\ & \text{Decreasing } \frac{\left(n_1 b_{12,1} h_1\right)^2}{\left(n_1 b_{22,1} h_1^2\right)} \\
\Rightarrow\ & \text{Increasing } \left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \frac{\left(n_1 b_{12,1} h_1\right)^2}{\left(n_1 b_{22,1} h_1^2\right)} \\
\Rightarrow\ & \text{Decreasing } \frac{\left(n_0 b_{12,0} h_0\right)^2}{\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \frac{\left(n_1 b_{12,1} h_1\right)^2}{\left(n_1 b_{22,1} h_1^2\right)}} \\
\Rightarrow\ & \text{Increasing } \left(n_0 b_{22,0} h_0^2\right) - \frac{\left(n_0 b_{12,0} h_0\right)^2}{\left(n_1 b_{11,1} + n_0 b_{11,0}\right) - \frac{\left(n_1 b_{12,1} h_1\right)^2}{\left(n_1 b_{22,1} h_1^2\right)}} \\
\Rightarrow\ & \text{Decreasing } V_{\beta_0}^D
\end{aligned}
$$

Thus the efficiency of the estimate β_0 increases as either specificity or sensitivity increases. Since the estimate of β_0 without diagnostic information included corresponds

to the case when sensitivity is the same as 1 - specificity ($p_0 = p_1$), the estimate 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, since $\frac{\partial h_i}{\partial \boldsymbol{\beta}}$ is the same for all subjects because $\boldsymbol{\beta} = (\beta_0)$, we can denote it as an unknown constant C_{β_0}. For $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)$, we have

$$\begin{aligned}
\frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} &= \frac{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1 + e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^2} \begin{bmatrix} 1 \\ z_{i,1} \end{bmatrix} \\
&= g_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} I(z_{i,1} = 0) + g_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} I(z_{i,1} = 1),
\end{aligned}$$

where g_0 and g_1 are the shorthand notation of $\frac{e^{\boldsymbol{\gamma}^T \mathbf{z}_i}}{\left(1+e^{\boldsymbol{\gamma}^T \mathbf{z}_i}\right)^2}$ when $z_{i,1} = 0$ and $z_{i,1} = 1$.

Let

$$\begin{aligned}
b_{11,0} &= E_{\mathbf{O}}\left(a_{11}^{(i)} + d_{11}^{(i)} \varphi_i(p_0, p_1) | z_{i,1} = 0\right), \\
b_{11,1} &= E_{\mathbf{O}}\left(a_{11}^{(i)} + d_{11}^{(i)} \varphi_i(p_0, p_1) | z_{i,1} = 1\right), \\
b_{12,0} &= E_{\mathbf{O}}\left(a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1) | z_{i,1} = 0\right), \\
b_{12,1} &= E_{\mathbf{O}}\left(a_{12}^{(i)} - d_{12}^{(i)} \varphi_i(p_0, p_1) | z_{i,1} = 1\right), \\
b_{22,0} &= E_{\mathbf{O}}\left(a_{22}^{(i)} + d_{22}^{(i)} \varphi_i(p_0, p_1) | z_{i,1} = 0\right), \\
b_{22,1} &= E_{\mathbf{O}}\left(a_{22}^{(i)} + d_{22}^{(i)} \varphi_i(p_0, p_1) | z_{i,1} = 1\right).
\end{aligned}$$

Also assume there are n_0 observations with $z_{i,1} = 0$, and n_1 observations with $z_{i,1} = 1$, by i.i.d. property when the covariates are the same, then

$$\begin{aligned}
E_{\mathbf{O}}\left(\mathbf{I}_{11}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n} \left[a_{11}^{(i)} + d_{11}^{(i)} \varphi_i(p_0, p_1)\right] \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}} \frac{\partial \pi_i}{\partial \boldsymbol{\gamma}^T}\right) \\
&= n_1 b_{11,1} g_1^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + n_0 b_{11,0} g_0^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} n_1 b_{11,1} g_1^2 + n_0 b_{11,0} g_0^2 & n_1 b_{11,1} g_1^2 \\ n_1 b_{11,1} g_1^2 & n_1 b_{11,1} g_1^2 \end{bmatrix}.
\end{aligned}$$

As a result,

$$E_{\mathbf{O}}^{-1}\left(\mathbf{I}_{11}\right) = \begin{bmatrix} \frac{1}{n_0 b_{11,0} g_0^2} & -\frac{1}{n_0 b_{11,0} g_0^2} \\ -\frac{1}{n_0 b_{11,0} g_0^2} & \frac{1}{n_0 b_{11,0} g_0^2} + \frac{1}{n_1 b_{11,1} g_1^2} \end{bmatrix}. \tag{4.37}$$

$$
\begin{aligned}
E_{\mathbf{O}}\left(\mathbf{I}_{12}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{12}^{(i)}-d_{12}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial\pi_i}{\partial\boldsymbol{\gamma}}\frac{\partial h_i}{\partial\boldsymbol{\beta}^T}\right)\\
&= C_{\beta_0}\left\{n_1 b_{12,1} g_1\begin{bmatrix}1\\1\end{bmatrix}+n_0 b_{12,0} g_0\begin{bmatrix}1\\0\end{bmatrix}\right\}\\
&= C_{\beta_0}\begin{bmatrix}n_1 b_{12,1} g_1+n_0 b_{12,0} g_0\\ n_1 b_{12,1} g_1\end{bmatrix}, \qquad (4.38)
\end{aligned}
$$

and

$$
\begin{aligned}
E_{\mathbf{O}}\left(\mathbf{I}_{22}\right) &= E_{\mathbf{O}}\left(\sum_{i=1}^{n}\left[a_{22}^{(i)}+d_{22}^{(i)}\varphi_i(p_0,p_1)\right]\frac{\partial h_i}{\partial\boldsymbol{\beta}}\frac{\partial h_i}{\partial\boldsymbol{\beta}^T}\right)\\
&= C_{\beta_0}^2\left(n_1 b_{22,1}+n_0 b_{22,0}\right). \qquad (4.39)
\end{aligned}
$$

From (4.37), (4.38), and (4.39),

$$
\begin{aligned}
\left[V_{\boldsymbol{\beta}}^D\right]^{-1} &= E_{\mathbf{O}}\left(\mathbf{I}_{22}\right)-E_{\mathbf{O}}^T\left(\mathbf{I}_{12}\right)E_{\mathbf{O}}^{-1}\left(\mathbf{I}_{11}\right)E_{\mathbf{O}}\left(\mathbf{I}_{12}\right)\\
&= C_{\beta_0}^2\left(n_1 b_{22,1}+n_0 b_{22,0}\right)-C_{\beta_0}^2\begin{bmatrix}n_1 b_{12,1} g_1+n_0 b_{12,0} g_0, & n_1 b_{12,1} g_1\end{bmatrix}\\
&\qquad\times\begin{bmatrix}\frac{1}{n_0 b_{11,0} g_0^2} & -\frac{1}{n_0 b_{11,0} g_0^2}\\ -\frac{1}{n_0 b_{11,0} g_0^2} & \frac{1}{n_0 b_{11,0} g_0^2}+\frac{1}{n_1 b_{11,1} g_1^2}\end{bmatrix}\begin{bmatrix}n_1 b_{12,1} g_1+n_0 b_{12,0} g_0\\ n_1 b_{12,1} g_1\end{bmatrix}\\
&= C_{\beta_0}^2\left[\left(n_1 b_{22,1}+n_0 b_{22,0}\right)-\frac{\left(n_0 b_{12,0} g_0\right)^2}{n_0 b_{11,0} g_0^2}-\frac{\left(n_1 b_{12,1} g_1\right)^2}{n_1 b_{11,1} g_1^2}\right]\\
&= C_{\beta_0}^2\left[\left(n_1 b_{22,1}+n_0 b_{22,0}\right)-\frac{n_0 b_{12,0}^2}{b_{11,0}}-\frac{n_1 b_{12,1}^2}{b_{11,1}}\right].
\end{aligned}
$$

Since $b_{11,1}$, $b_{11,0}$, $b_{22,0}$, $b_{22,1}$ are increasing functions of p_0, and decreasing functions of p_1 through their dependence of $\varphi_i(p_0,p_1)$, and $b_{12,1}$, $b_{12,0}$ are decreasing functions of p_0, and increasing functions of p_1 through their dependence of $\varphi_i(p_0,p_1)$, $\left[V_{\boldsymbol{\beta}}^D\right]^{-1}$ is an increasing function of p_0, and decreasing functions of p_1. Larger $\left[V_{\boldsymbol{\beta}}^D\right]^{-1}$ lead to smaller $V_{\boldsymbol{\beta}}^D$. Thus the efficiency of the estimate $\hat{\boldsymbol{\beta}}$ increases as either specificity or sensitivity increases, , and the estimate of $\boldsymbol{\beta}$ with diagnostic information included is more efficient than that without diagnostic information included.

For $V_{\boldsymbol{\gamma}}^D$, it can be computed as

$$
\frac{1}{C_{\beta_0}^2 G}\begin{bmatrix}V_{11} & V_{12}\\ V_{12} & V_{22}\end{bmatrix},
$$

where

$$
\begin{aligned}
G &= \frac{1}{\left(n_1 b_{11,1} g_1^2\right)\left(n_0 b_{11,0} g_0^2\right)\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \left(n_1 b_{11,1} g_1^2\right)\left(n_0 b_{12,0} g_0\right)^2 - \left(n_0 b_{11,0} g_0^2\right)\left(n_1 b_{12,1} g_1\right)^2}, \\
V_{11} &= \left(n_1 b_{11,1} g_1^2\right)\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \left(n_1 b_{12,1} g_1\right)^2, \\
V_{22} &= \left(n_1 b_{11,1} g_1^2 + n_0 b_{11,0} g_0^2\right)\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \left(n_1 b_{12,1} g_1 + n_0 b_{12,0} g_0\right)^2, \\
V_{12} &= -\left(n_1 b_{22,1} g_1^2\right)\left(n_1 b_{11,1} + n_0 b_{11,0}\right) + \left(n_1 b_{12,1} g_1 + n_0 b_{12,0} g_0\right)\left(n_1 b_{12,1} g_1\right).
\end{aligned}
$$

For the variance of $\hat{\gamma}_0$, we have

$$
\begin{aligned}
& C_{\beta_0}^2 V_{\gamma_0}^D \\
= & \frac{\left(n_1 b_{11,1} g_1^2\right)\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \left(n_1 b_{12,1} g_1\right)^2}{\left(n_1 b_{11,1} g_1^2\right)\left(n_0 b_{11,0} g_0^2\right)\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \left(n_1 b_{11,1} g_1^2\right)\left(n_0 b_{12,0} g_0\right)^2 - \left(n_0 b_{11,0} g_0^2\right)\left(n_1 b_{12,1} g_1\right)^2} \\
= & \frac{1}{\left(n_0 b_{11,0} g_0^2\right) - \frac{\left(n_1 b_{11,1} g_1^2\right)\left(n_0 b_{12,0} g_0\right)^2}{\left(n_1 b_{11,1} g_1^2\right)\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \left(n_1 b_{12,1} g_1\right)^2}} \\
= & \frac{1}{\left(n_0 b_{11,0} g_0^2\right) - \frac{\left(n_0 b_{12,0} g_0\right)^2}{\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \frac{\left(n_1 b_{12,1} g_1\right)^2}{\left(n_1 b_{11,1} g_1^2\right)}}}
\end{aligned}
$$

Because $b_{11,1}$, $b_{11,0}$, $b_{22,0}$, $b_{22,1}$ are increasing functions of p_0, and decreasing functions of p_1 through their dependence of $\varphi_i(p_0, p_1)$, and $b_{12,1}$, $b_{12,0}$ are decreasing functions of p_0, and increasing functions of p_1 through their dependence of $\varphi_i(p_0, p_1)$, we can infer that

$$
\begin{aligned}
& \text{Increasing } p_0\text{, and/or decreasing } p_1 \\
\Rightarrow\ & \text{Increasing } b_{11,1}, b_{11,0}, b_{22,0}, b_{22,1}\text{, and decreasing } b_{12,1}, b_{12,0} \\
\Rightarrow\ & \text{Decreasing } \frac{\left(n_1 b_{12,1} g_1\right)^2}{\left(n_1 b_{11,1} g_1^2\right)} \\
\Rightarrow\ & \text{Increasing } \left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \frac{\left(n_1 b_{12,1} g_1\right)^2}{\left(n_1 b_{11,1} g_1^2\right)} \\
\Rightarrow\ & \text{Decreasing } \frac{\left(n_0 b_{12,0} g_0\right)^2}{\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \frac{\left(n_1 b_{12,1} g_1\right)^2}{\left(n_1 b_{11,1} g_1^2\right)}} \\
\Rightarrow\ & \text{Increasing } \left(n_0 b_{11,0} g_0^2\right) - \frac{\left(n_0 b_{12,0} g_0\right)^2}{\left(n_1 b_{22,1} + n_0 b_{22,0}\right) - \frac{\left(n_1 b_{12,1} g_1\right)^2}{\left(n_1 b_{11,1} g_1^2\right)}} \\
\Rightarrow\ & \text{Decreasing } V_{\gamma_0}^D
\end{aligned}
$$

Thus the efficiency of the estimate γ_0 increases as either specificity or sensitivity increases. Since the estimate of γ_0 without diagnostic information included corresponds

to the case when sensitivity is the same as 1 - specificity ($p_0 = p_1$), the estimate of γ_0 with diagnostic information included (with $p_0 > p_1$) is more efficient than that without diagnostic information included.

□

4.2 General Simulation Setup

With additional sensitivity and specificity information incorporated in the model, it is expected that more efficient estimators can be obtained. Section 4.1 also showed, under the logistic-exponential mixture, we are expected to see the gain in the relative efficiency. Furthermore, we will compare the following three models through simulation studies:

1. Additional diagnostic information is not included in the model.
2. Additional diagnostic information is included in the model. Sensitivity and specificity parameters will be estimated using the procedure covered in the previous section.
3. Additional diagnostic information is included in the model. Sensitivity and specificity parameters are known a priori.

Cure rate models with semi-parametric PH or semi-parametric AFT specifications are considered separately. For baseline hazard function, Weibull distribution will be used. The PDF and hazard function of Weibull distribution are

$$f_0(t|k,h) = kh(ht)^{k-1}e^{-(ht)^k},$$

and

$$h_0(t|k,h) = kh(ht)^{k-1},$$

respectively, where k is the shape parameter and h is the scale parameter.

Four different sets of Weibull distribution are considered: 1) $h = 1$, $k = 2$; 2) $h = 2$, $k = 2$; 3) $h = \frac{2}{3}$, $k = 3$; 4) $h = \frac{1}{3}$, $k = 4$. The shape of these four density functions are shown in Figure 4.1.

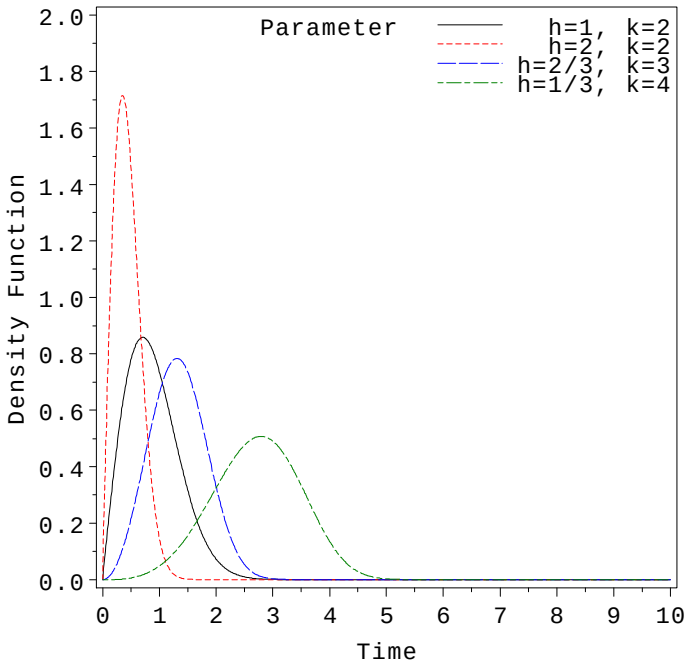


Figure 4.1: Weibull Density Curves

The same covariates are used in both the “incidence” part and the “latency” part of the model. Specification for the "latency" part may be semi-parametric PH model, or semi-parametric AFT model. The covariates include a three-level treatment effect and a two-level gender effect. All these true parameters are maintained the same throughout the simulation studies. Different sensitivity and specificity combination scenarios are used in the simulations.

Maximum follow-up time is fixed at 6 in the simulation, with uniform censoring distribution throughout the follow-up period. A subject is censored ($\delta_i = 0$) if the simulated censoring time is before the simulated event time. If we denote y as the censoring time, and x as the event time, we can compute the probability of censoring

as:

$$\begin{aligned} P(y < x) &= \int_0^6 \frac{1}{6} dy \int_y^6 kh(hx)^{k-1} e^{-(hx)^k} dx \\ &= \int_0^6 \frac{1}{6} \left[e^{-(hy)^k} - e^{-(6h)^k} \right] dy. \end{aligned}$$

So, for each of the four different baseline hazard functions that we are considering, the expected censoring rates are 1) 0.1477 when $h = 1$, $k = 2$; 2) 0.0739 when $h = 2$, $k = 2$; 3) 0.2232 when $h = \frac{2}{3}$, $k = 3$; 4) 0.4532 when $h = \frac{1}{3}$, $k = 4$. The expected censoring rates for each simulated observations are different, depending on whether the AFT or PH model is used, and depending on the simulated covariates. When the censoring rates are higher, higher efficiency gains are expected since the diagnostic information will help to reduce the ambiguity for more censored observations.

The incidence part is modeled using logit link with the covariates specified as:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \gamma_0 + \gamma_1 I_{(\text{TRT=1})} + \gamma_2 I_{(\text{TRT=2})} + \gamma_3 I_{(\text{SEX=Male})} \tag{4.40}$$

Cured status c_i is simulated based on π_i. After censoring status is determined and c_i is simulated, d_i is simulated based on the following Bernoulli trial:

$$d_i | \, (c_i = 0, \delta_i = 0) \sim Bernoulli(p_0),$$

and

$$d_i | \, (c_i = 1, \delta_i = 0) \sim Bernoulli(p_1).$$

For each run, two hundreds subjects are simulated. A total of 1000 simulated runs is performed for each simulation scenario.

4.3 Evaluation of Semi-Parametric PH Cure Rate Models

Analyses of the simulated data are performed using the model as specified in the previous section. The latency part is modeled using semi-parametric PH model. The covariates for this PH are specified as:

$$h_u(t_i) = h_0(t_i) \exp\left(\beta_1 I_{(\text{TRT=1})} + \beta_2 I_{(\text{TRT=2})} + \beta_3 I_{(\text{SEX=Male})}\right). \tag{4.41}$$

For all the cases presented in the general setup section, simulations with 100% subjects with available diagnostic information were performed. To investigate the impact of the estimation when some subjects do not have available diagnostic information, simulations are also performed for Weibull distribution with $h = \frac{1}{3}$ and $k = 4$, and with 50% subjects without diagnostic information. In the following sub-sections, the simulation results for each of the four different baseline hazard functions are presented. A cross-comparison of all the simulation results is presented at the end of this section.

4.3.1 Models using Weibull distribution with parameter $h = 1$ and $k = 2$

For models using Weibull distribution with parameter $h = 1$ and $k = 2$, the results with 80% specificity ($p_1 = 0.2$) are shown in Appendix Tables A.1 (assuming known sensitivity and specificity) and A.2 (assuming unknown sensitivity and specificity); the results with 90% specificity ($p_1 = 0.1$) are shown in Appendix Tables A.3(assuming known sensitivity and specificity) and A.4 (assuming unknown sensitivity and specificity); and the results with 100% specificity ($p_1 = 0$) are shown in Tables 4.1 (assuming known sensitivity and specificity) and 4.2 (assuming unknown sensitivity and specificity). Corresponding figures for the relative efficiency, MSE, and bias are presented from Appendix Figures A.1 through A.4, and from Figures 4.2 through 4.3.

Under each of the simulation scenarios, sensitivity is assessed under 5 different scenarios: 0, 0.3, 0.5, 0.7, 1. When sensitivity is 0, the model without the use of additional diagnostic information is applied. In each of the plots, we can see that the biases are always maintained within a close range of true values, which indicates the asymptotic consistency of the proposed estimators. Within each plot or table, the relative efficiency is increasing with the increase of sensitivity.

When the sensitivity and specificity are known, if we look across plots (Figures A.1, A.3, and 4.2) or tables (Tables A.1, A.3, and 4.1), we can see for the same parameter, the efficiency increases with the increase of specificity. For example, for γ_0, when specificity is 80% and sensitivity is 100%, the relative efficiency is 1.1866. This relative efficiency increases to 1.2041 when the sensitivity is 90%, and to 1.2528 when the sensitivity is

100%. Similar observations can also be make when the sensitivity and specificity are unknown (Figures A.2, A.4, and 4.3; Tables A.2, A.4, and 4.2). The extent of relative efficiency gain is less when the sensitivity and specificity are unknown, compared to when the sensitivity and specificity are known. For example, when both the sensitivity and specificity are 100%, if these parameters are known, the relative efficiency gain for γ_0 is 1.2528; if these parameters are unknown, the relative efficiency gain for γ_0 is 1.1823.

Table 4.1: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1870/-0.0130	0.1883/-0.0117	0.1874/-0.0126	0.1898/-0.0102	0.1897/-0.0103
	SD		0.2919	0.2905	0.2906	0.2883	0.2800
	MSE		0.0854	0.0845	0.0846	0.0832	0.0785
	Relative Efficiency		**1.0000**	**1.0093**	**1.0087**	**1.0245**	**1.0866**
β_2	Mean/Bias	-0.3	-0.3234/-0.0234	-0.3222/-0.0222	-0.3212/-0.0212	-0.3183/-0.0183	-0.3172/-0.0172
	SD		0.2859	0.2833	0.2822	0.2805	0.2664
	MSE		0.0823	0.0807	0.0801	0.0790	0.0713
	Relative Efficiency		**1.0000**	**1.0187**	**1.0266**	**1.0386**	**1.1516**
β_3	Mean/Bias	0.1	0.1171/0.0171	0.1168/0.0168	0.1170/0.0170	0.1161/0.0161	0.1133/0.0133
	SD		0.2312	0.2301	0.2286	0.2274	0.2213
	MSE		0.0538	0.0532	0.0525	0.0520	0.0491
	Relative Efficiency		**1.0000**	**1.0094**	**1.0235**	**1.0341**	**1.0921**
γ_0	Mean/Bias	0.25	0.2619/0.0119	0.2641/0.0141	0.2584/0.0084	0.2640/0.0140	0.2558/0.0058
	SD		0.3335	0.3288	0.3263	0.3184	0.2979
	MSE		0.1113	0.1083	0.1065	0.1016	0.0888
	Relative Efficiency		**1.0000**	**1.0288**	**1.0443**	**1.0966**	**1.2528**
γ_1	Mean/Bias	-0.1	-0.1022/-0.0022	-0.1049/-0.0049	-0.1023/-0.0023	-0.1083/-0.0083	-0.1034/-0.0034
	SD		0.4004	0.3945	0.3847	0.3769	0.3558
	MSE		0.1603	0.1556	0.1480	0.1421	0.1266
	Relative Efficiency		**1.0000**	**1.0301**	**1.0831**	**1.1283**	**1.2661**
γ_2	Mean/Bias	0.5	0.5127/0.0127	0.5066/0.0066	0.5055/0.0055	0.4985/-0.0015	0.5007/0.0007
	SD		0.4297	0.4190	0.4133	0.3975	0.3654
	MSE		0.1848	0.1756	0.1709	0.1580	0.1335
	Relative Efficiency		**1.0000**	**1.0520**	**1.0809**	**1.1685**	**1.3830**
γ_3	Mean/Bias	-0.1	-0.1098/-0.0098	-0.1113/-0.0113	-0.1074/-0.0074	-0.1070/-0.0070	-0.1037/-0.0037
	SD		0.3382	0.3314	0.3269	0.3149	0.3006
	MSE		0.1145	0.1100	0.1069	0.0992	0.0904
	Relative Efficiency		**1.0000**	**1.0415**	**1.0701**	**1.1533**	**1.2660**

Figure 4.2: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity

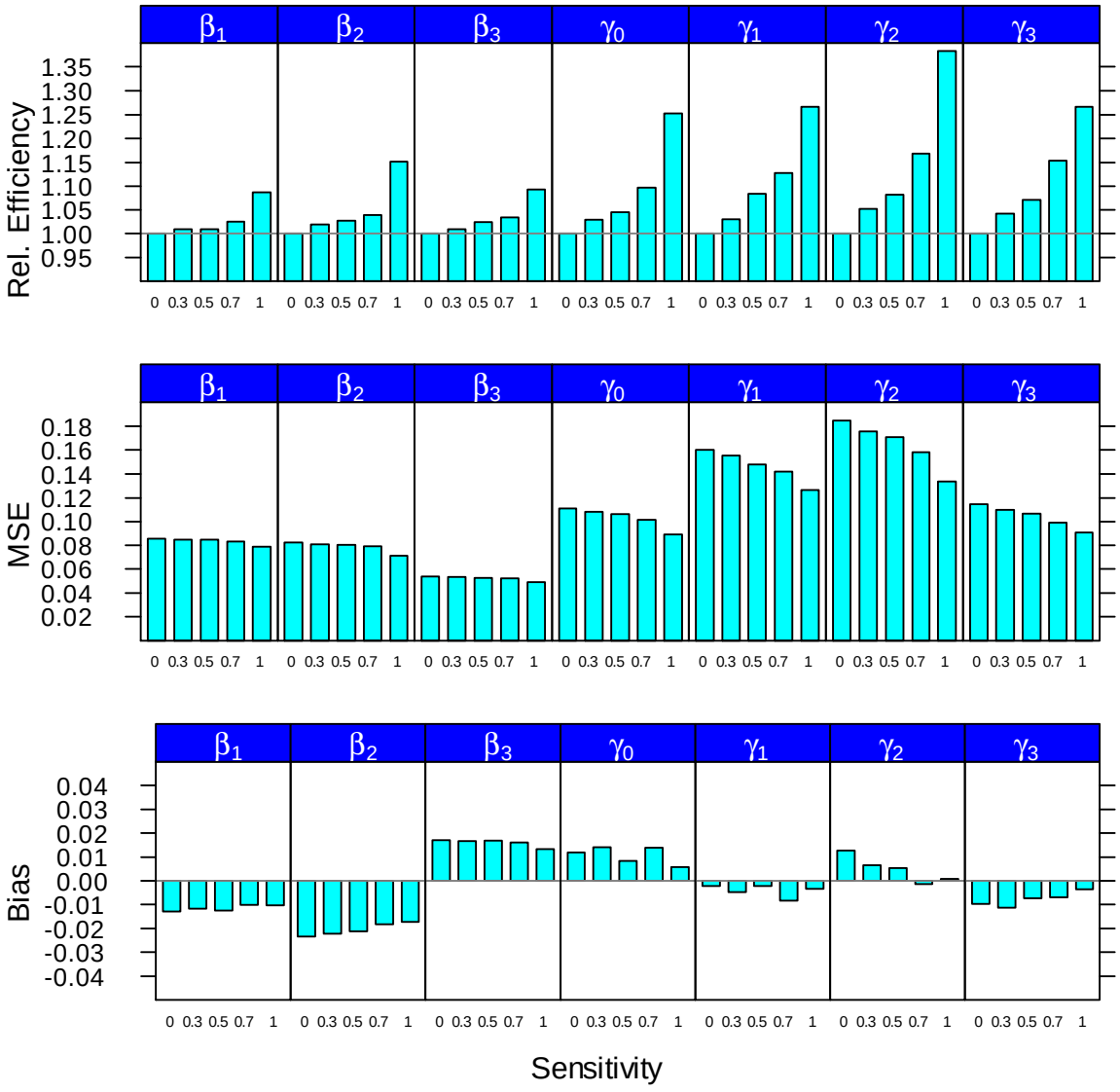
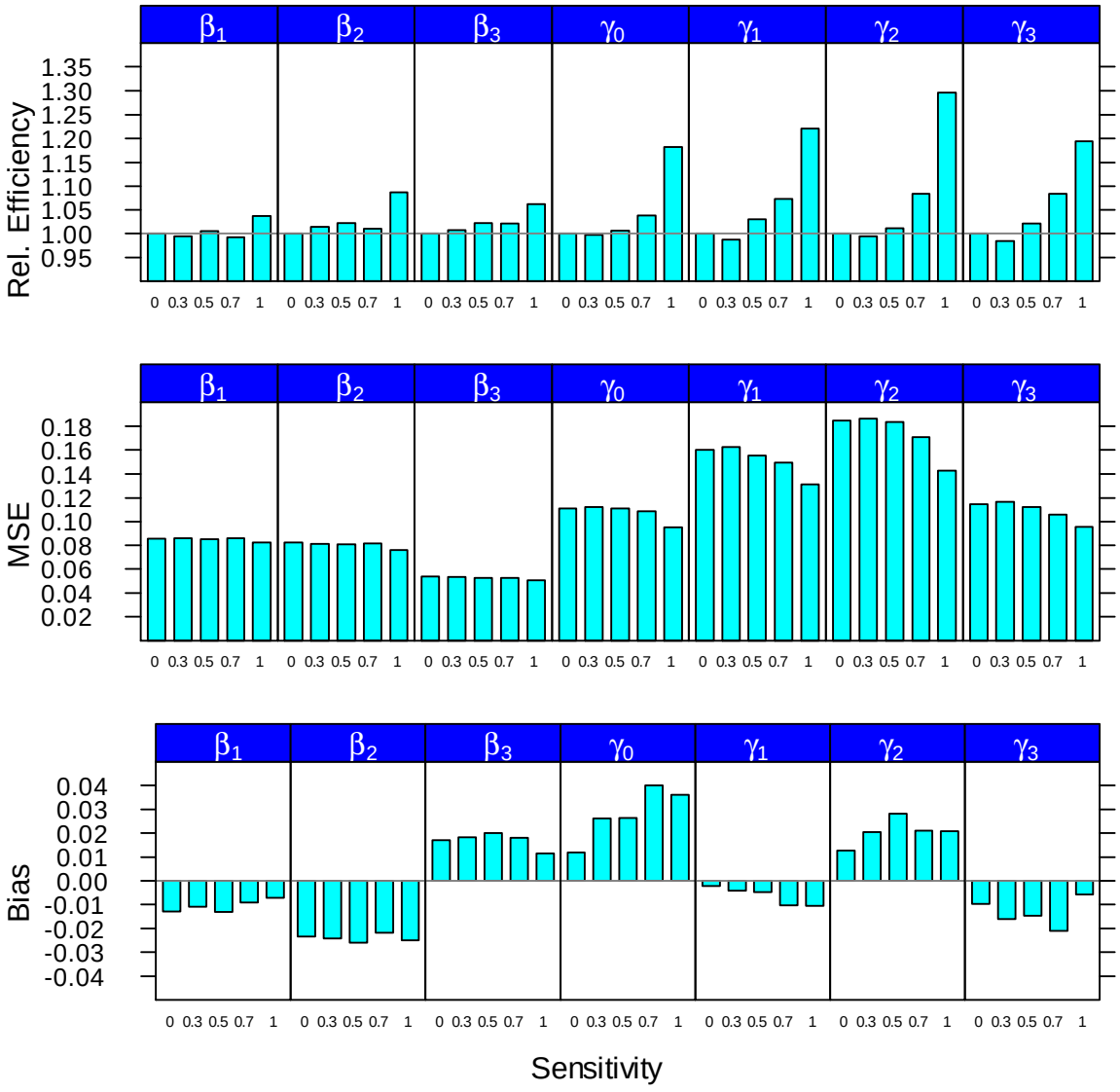


Table 4.2: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1870/-0.0130	0.1890/-0.0110	0.1868/-0.0132	0.1909/-0.0091	0.1929/-0.0071
	SD		0.2919	0.2927	0.2912	0.2929	0.2867
	MSE		0.0854	0.0858	0.0850	0.0859	0.0822
	Relative Efficiency		**1.0000**	**0.9943**	**1.0047**	**0.9927**	**1.0365**
β_2	Mean/Bias	-0.3	-0.3234/-0.0234	-0.3242/-0.0242	-0.3259/-0.0259	-0.3219/-0.0219	-0.3249/-0.0249
	SD		0.2859	0.2840	0.2829	0.2845	0.2743
	MSE		0.0823	0.0812	0.0807	0.0814	0.0759
	Relative Efficiency		**1.0000**	**1.0137**	**1.0216**	**1.0101**	**1.0864**
β_3	Mean/Bias	0.1	0.1171/0.0171	0.1184/0.0184	0.1200/0.0200	0.1181/0.0181	0.1116/0.0116
	SD		0.2312	0.2304	0.2287	0.2289	0.2244
	MSE		0.0538	0.0534	0.0527	0.0527	0.0505
	Relative Efficiency		**1.0000**	**1.0071**	**1.0218**	**1.0206**	**1.0618**
γ_0	Mean/Bias	0.25	0.2619/0.0119	0.2762/0.0262	0.2764/0.0264	0.2901/0.0401	0.2861/0.0361
	SD		0.3335	0.3340	0.3324	0.3274	0.3067
	MSE		0.1113	0.1122	0.1112	0.1088	0.0954
	Relative Efficiency		**1.0000**	**0.9969**	**1.0065**	**1.0375**	**1.1823**
γ_1	Mean/Bias	-0.1	-0.1022/-0.0022	-0.1042/-0.0042	-0.1048/-0.0048	-0.1103/-0.0103	-0.1106/-0.0106
	SD		0.4004	0.4029	0.3945	0.3865	0.3623
	MSE		0.1603	0.1624	0.1556	0.1495	0.1314
	Relative Efficiency		**1.0000**	**0.9872**	**1.0299**	**1.0728**	**1.2210**
γ_2	Mean/Bias	0.5	0.5127/0.0127	0.5205/0.0205	0.5282/0.0282	0.5211/0.0211	0.5208/0.0208
	SD		0.4297	0.4310	0.4274	0.4129	0.3774
	MSE		0.1848	0.1862	0.1834	0.1710	0.1429
	Relative Efficiency		**1.0000**	**0.9942**	**1.0111**	**1.0830**	**1.2962**
γ_3	Mean/Bias	-0.1	-0.1098/-0.0098	-0.1160/-0.0160	-0.1148/-0.0148	-0.1211/-0.0211	-0.1059/-0.0059
	SD		0.3382	0.3409	0.3348	0.3249	0.3095
	MSE		0.1145	0.1165	0.1123	0.1060	0.0958
	Relative Efficiency		**1.0000**	**0.9843**	**1.0208**	**1.0837**	**1.1941**

Figure 4.3: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity



4.3.2 Models using Weibull distribution with parameter $h = 2$ and $k = 2$

For models using Weibull distribution with parameter $h = 2$ and $k = 2$, the results with 80% specificity ($p_1 = 0.2$) are shown in Appendix Tables A.5 (assuming known sensitivity and specificity) and A.6 (assuming unknown sensitivity and specificity); the results with 90% specificity ($p_1 = 0.1$) are shown in Appendix Tables A.7(assuming known sensitivity and specificity) and A.8 (assuming unknown sensitivity and specificity); and the results with 100% specificity ($p_1 = 0$) are shown in Tables 4.3 (assuming known sensitivity and specificity) and 4.4 (assuming unknown sensitivity and specificity). Corresponding figures for the relative efficiency, MSE, and bias are presented from Appendix Figures A.5 through A.8, and from Figures 4.4 through 4.5.

Similar observations can be made as for the case when the baseline Weibull distribution with parameter $h = 1$ and $k = 2$. Biases are small when comparing to true values. The relative efficiency can be seen to increase with higher diagnostic sensitivity and specificity. This efficiency gain is higher when the sensitivity and specificity information are known. The extent of such increase, is relatively smaller compared to the case when the baseline Weibull distribution with parameter $h = 1$ and $k = 2$. For example, when both the sensitivity and specificity are known to be 100%, the relative efficiency gain for γ_0 is 1.2528 when $h = 1$ and $k = 2$, while the relative efficiency gain for γ_0 under the same simulation scenario when $h = 2$ and $k = 2$ is 1.0896.

Table 4.3: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2053/0.0053	0.2054/0.0054	0.2063/0.0063	0.2050/0.0050	0.2061/0.0061
	SD		0.2733	0.2735	0.2733	0.2718	0.2649
	MSE		0.0747	0.0748	0.0747	0.0739	0.0702
	Relative Efficiency		**1.0000**	**0.9988**	**1.0003**	**1.0115**	**1.0647**
β_2	Mean/Bias	-0.3	-0.3036/-0.0036	-0.3032/-0.0032	-0.3035/-0.0035	-0.3039/-0.0039	-0.3026/-0.0026
	SD		0.2586	0.2582	0.2581	0.2568	0.2485
	MSE		0.0669	0.0667	0.0666	0.0659	0.0618
	Relative Efficiency		**1.0000**	**1.0033**	**1.0036**	**1.0143**	**1.0825**
β_3	Mean/Bias	0.1	0.0997/-0.0003	0.0992/-0.0008	0.0997/-0.0003	0.0993/-0.0007	0.0983/-0.0017
	SD		0.2188	0.2186	0.2186	0.2185	0.2127
	MSE		0.0479	0.0478	0.0478	0.0477	0.0452
	Relative Efficiency		**1.0000**	**1.0014**	**1.0014**	**1.0027**	**1.0577**
γ_0	Mean/Bias	0.25	0.2495/-0.0005	0.2487/-0.0013	0.2517/0.0017	0.2506/0.0006	0.2510/0.0010
	SD		0.3099	0.3072	0.3050	0.3039	0.2969
	MSE		0.0961	0.0944	0.0930	0.0923	0.0882
	Relative Efficiency		**1.0000**	**1.0176**	**1.0326**	**1.0404**	**1.0896**
γ_1	Mean/Bias	-0.1	-0.0860/0.0140	-0.0861/0.0139	-0.0895/0.0105	-0.0867/0.0133	-0.0903/0.0097
	SD		0.3696	0.3678	0.3628	0.3606	0.3537
	MSE		0.1368	0.1355	0.1317	0.1302	0.1252
	Relative Efficiency		**1.0000**	**1.0098**	**1.0384**	**1.0511**	**1.0925**
γ_2	Mean/Bias	0.5	0.5138/0.0138	0.5136/0.0136	0.5109/0.0109	0.5122/0.0122	0.5085/0.0085
	SD		0.3827	0.3795	0.3739	0.3711	0.3644
	MSE		0.1466	0.1442	0.1400	0.1379	0.1329
	Relative Efficiency		**1.0000**	**1.0165**	**1.0472**	**1.0631**	**1.1028**
γ_3	Mean/Bias	-0.1	-0.1097/-0.0097	-0.1086/-0.0086	-0.1088/-0.0088	-0.1084/-0.0084	-0.1053/-0.0053
	SD		0.3271	0.3237	0.3211	0.3207	0.3103
	MSE		0.1071	0.1048	0.1032	0.1029	0.0963
	Relative Efficiency		**1.0000**	**1.0213**	**1.0376**	**1.0402**	**1.1114**

Figure 4.4: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity

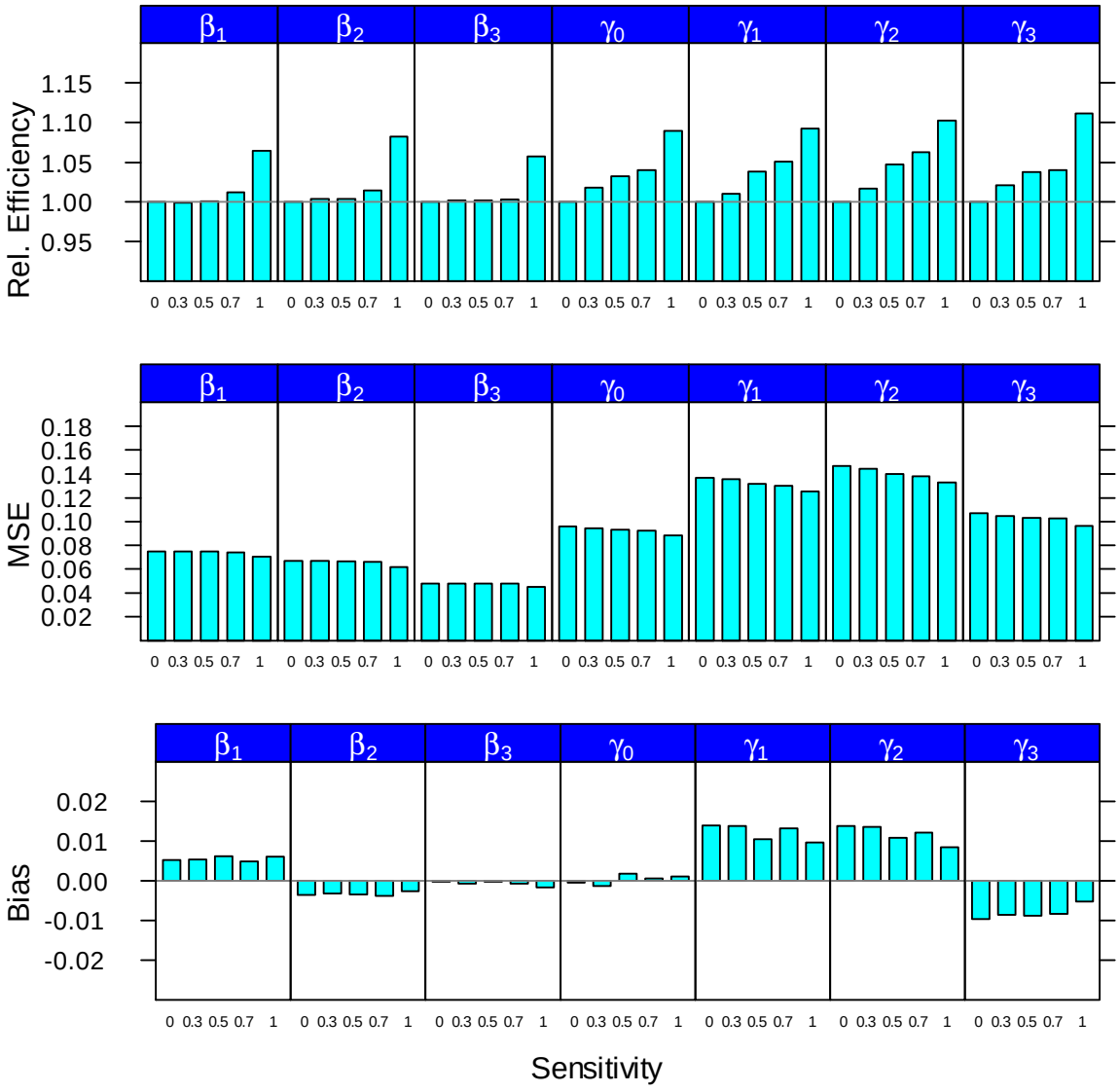
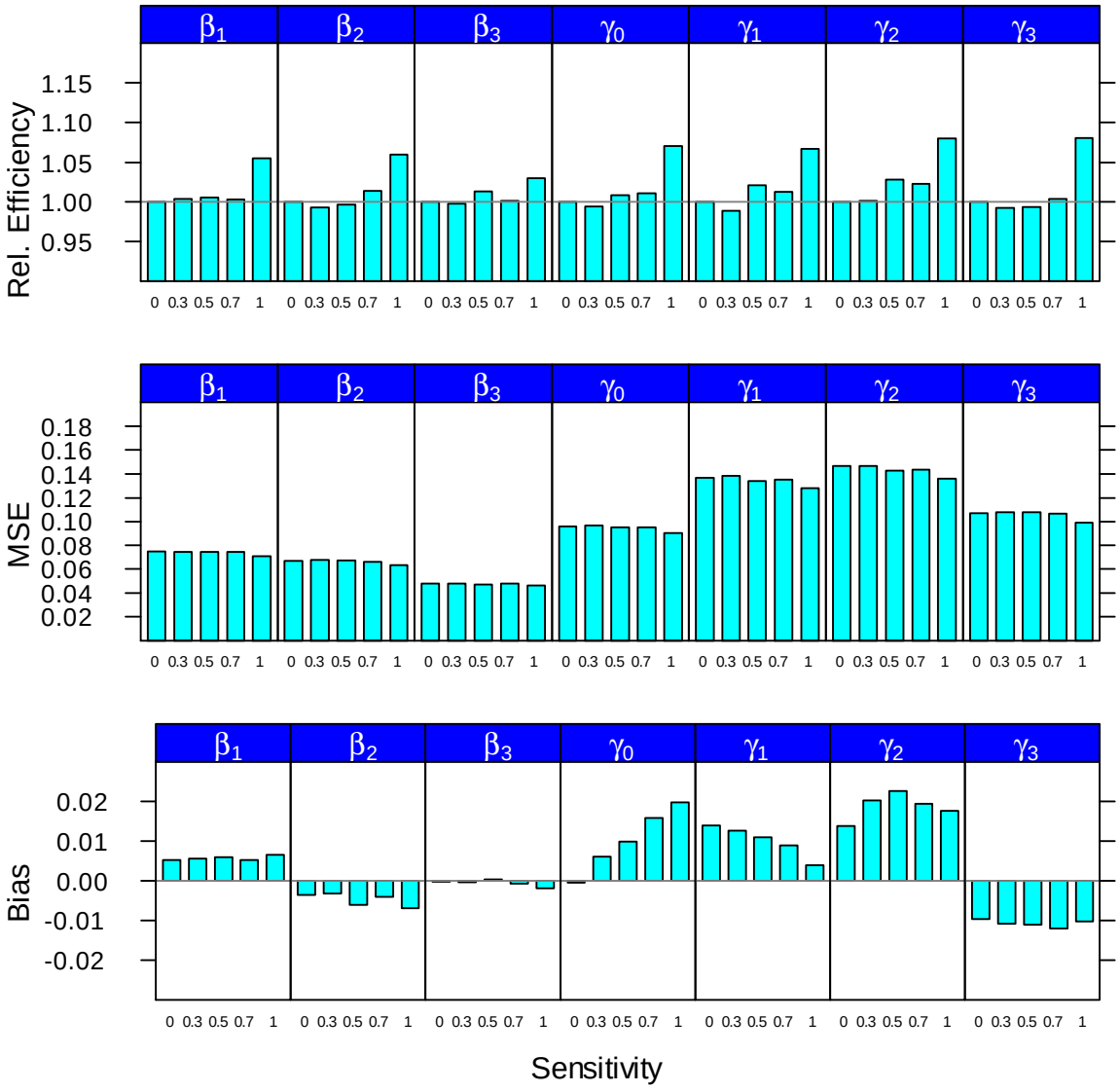


Table 4.4: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2053/0.0053	0.2056/0.0056	0.2060/0.0060	0.2053/0.0053	0.2066/0.0066
	SD		0.2733	0.2729	0.2726	0.2730	0.2661
	MSE		0.0747	0.0745	0.0744	0.0745	0.0709
	Relative Efficiency		**1.0000**	**1.0034**	**1.0053**	**1.0029**	**1.0550**
β_2	Mean/Bias	-0.3	-0.3036/-0.0036	-0.3033/-0.0033	-0.3061/-0.0061	-0.3041/-0.0041	-0.3069/-0.0069
	SD		0.2586	0.2595	0.2591	0.2568	0.2512
	MSE		0.0669	0.0674	0.0672	0.0660	0.0631
	Relative Efficiency		**1.0000**	**0.9927**	**0.9963**	**1.0138**	**1.0600**
β_3	Mean/Bias	0.1	0.0997/-0.0003	0.0996/-0.0004	0.1003/0.0003	0.0993/-0.0007	0.0981/-0.0019
	SD		0.2188	0.2190	0.2173	0.2186	0.2155
	MSE		0.0479	0.0480	0.0472	0.0478	0.0464
	Relative Efficiency		**1.0000**	**0.9976**	**1.0130**	**1.0010**	**1.0304**
γ_0	Mean/Bias	0.25	0.2495/-0.0005	0.2561/0.0061	0.2599/0.0099	0.2659/0.0159	0.2698/0.0198
	SD		0.3099	0.3109	0.3087	0.3083	0.2995
	MSE		0.0961	0.0967	0.0954	0.0953	0.0901
	Relative Efficiency		**1.0000**	**0.9939**	**1.0082**	**1.0105**	**1.0708**
γ_1	Mean/Bias	-0.1	-0.0860/0.0140	-0.0873/0.0127	-0.0890/0.0110	-0.0910/0.0090	-0.0960/0.0040
	SD		0.3696	0.3718	0.3658	0.3674	0.3579
	MSE		0.1368	0.1384	0.1339	0.1351	0.1281
	Relative Efficiency		**1.0000**	**0.9885**	**1.0210**	**1.0123**	**1.0669**
γ_2	Mean/Bias	0.5	0.5138/0.0138	0.5202/0.0202	0.5226/0.0226	0.5194/0.0194	0.5177/0.0177
	SD		0.3827	0.3825	0.3774	0.3784	0.3682
	MSE		0.1466	0.1467	0.1429	0.1435	0.1359
	Relative Efficiency		**1.0000**	**1.0008**	**1.0281**	**1.0229**	**1.0801**
γ_3	Mean/Bias	-0.1	-0.1097/-0.0097	-0.1108/-0.0108	-0.1111/-0.0111	-0.1120/-0.0120	-0.1102/-0.0102
	SD		0.3271	0.3284	0.3282	0.3266	0.3146
	MSE		0.1071	0.1080	0.1078	0.1068	0.0991
	Relative Efficiency		**1.0000**	**0.9919**	**0.9933**	**1.0034**	**1.0808**

Figure 4.5: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity



4.3.3 Models using Weibull distribution with parameter $h = \frac{2}{3}$ and $k = 3$

For models using Weibull distribution with parameter $h = \frac{2}{3}$ and $k = 3$, the results with 80% specificity ($p_1 = 0.2$) are shown in Appendix Tables A.9 (assuming known sensitivity and specificity) and A.10 (assuming unknown sensitivity and specificity); the results with 90% specificity ($p_1 = 0.1$) are shown in Appendix Tables A.11(assuming known sensitivity and specificity) and A.12 (assuming unknown sensitivity and specificity); and the results with 100% specificity ($p_1 = 0$) are shown in Tables 4.5 (assuming known sensitivity and specificity) and 4.6 (assuming unknown sensitivity and specificity). Corresponding figures for the relative efficiency, MSE, and bias are presented from Appendix Figures A.9 through A.12, and from Figures 4.6 through 4.7.

Similar observations can be made as for the case when the baseline Weibull distribution with parameter $h = 1$ and $k = 2$. Biases are small when comparing to true values. The relative efficiency can be seen to increase with higher diagnostic sensitivity and specificity. This efficiency gain is higher when the sensitivity and specificity information are known. The extent of such increase, is larger compared to the case when the baseline Weibull distribution with parameter $h = 1$ and $k = 2$. For example, when both the sensitivity and specificity are known to be 100%, the relative efficiency gain for γ_0 is 1.2528 when $h = 1$ and $k = 2$, while the relative efficiency gain for γ_0 under the same simulation scenario when $h = \frac{2}{3}$ and $k = 3$ is 1.4777.

Table 4.5: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1985/-0.0015	0.1966/-0.0034	0.1981/-0.0019	0.1970/-0.0030	0.1995/-0.0005
	SD		0.3258	0.3239	0.3219	0.3207	0.3048
	MSE		0.1061	0.1049	0.1036	0.1029	0.0929
	Relative Efficiency		**1.0000**	**1.0116**	**1.0243**	**1.0319**	**1.1419**
β_2	Mean/Bias	-0.3	-0.3224/-0.0224	-0.3220/-0.0220	-0.3231/-0.0231	-0.3198/-0.0198	-0.3176/-0.0176
	SD		0.3085	0.3053	0.3054	0.3031	0.2915
	MSE		0.0957	0.0937	0.0938	0.0923	0.0853
	Relative Efficiency		**1.0000**	**1.0211**	**1.0199**	**1.0356**	**1.1198**
β_3	Mean/Bias	0.1	0.0953/-0.0047	0.0944/-0.0056	0.0942/-0.0058	0.0963/-0.0037	0.0940/-0.0060
	SD		0.2421	0.2408	0.2388	0.2371	0.2259
	MSE		0.0586	0.0580	0.0571	0.0562	0.0511
	Relative Efficiency		**1.0000**	**1.0108**	**1.0273**	**1.0422**	**1.1485**
γ_0	Mean/Bias	0.25	0.2608/0.0108	0.2605/0.0105	0.2590/0.0090	0.2611/0.0111	0.2611/0.0111
	SD		0.3560	0.3446	0.3303	0.3211	0.2929
	MSE		0.1269	0.1189	0.1092	0.1032	0.0859
	Relative Efficiency		**1.0000**	**1.0673**	**1.1619**	**1.2297**	**1.4777**
γ_1	Mean/Bias	-0.1	-0.1120/-0.0120	-0.1077/-0.0077	-0.1100/-0.0100	-0.1083/-0.0083	-0.1074/-0.0074
	SD		0.4304	0.4216	0.4097	0.3928	0.3712
	MSE		0.1854	0.1778	0.1680	0.1544	0.1378
	Relative Efficiency		**1.0000**	**1.0420**	**1.1034**	**1.2003**	**1.3444**
γ_2	Mean/Bias	0.5	0.5175/0.0175	0.5141/0.0141	0.5142/0.0142	0.5060/0.0060	0.5042/0.0042
	SD		0.4424	0.4270	0.4148	0.4074	0.3775
	MSE		0.1960	0.1825	0.1722	0.1660	0.1426
	Relative Efficiency		**1.0000**	**1.0732**	**1.1375**	**1.1792**	**1.3728**
γ_3	Mean/Bias	-0.1	-0.0949/0.0051	-0.0934/0.0066	-0.0928/0.0072	-0.0945/0.0055	-0.0919/0.0081
	SD		0.3652	0.3536	0.3418	0.3247	0.3053
	MSE		0.1334	0.1251	0.1169	0.1054	0.0933
	Relative Efficiency		**1.0000**	**1.0666**	**1.1416**	**1.2654**	**1.4307**

Figure 4.6: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity

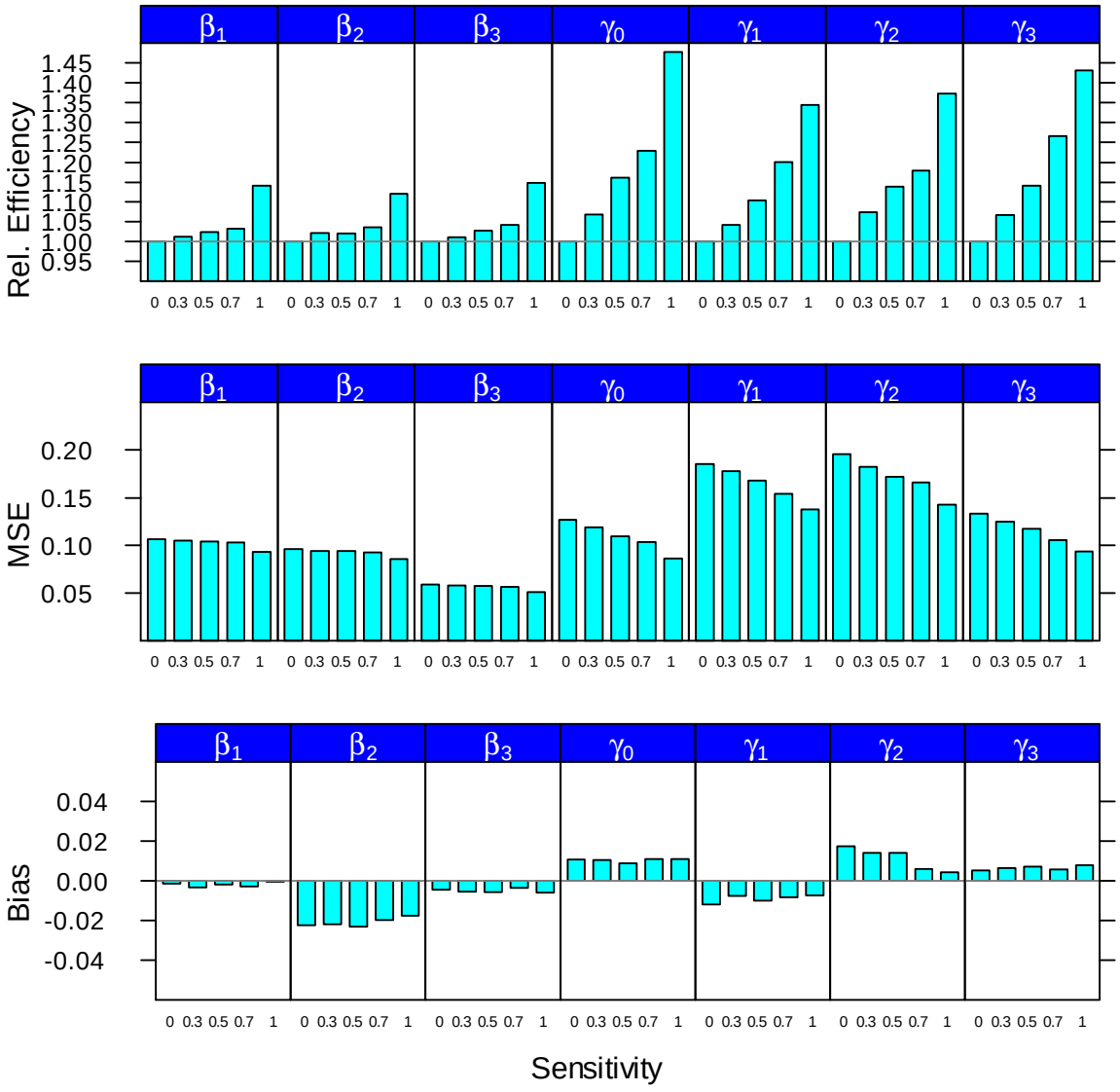
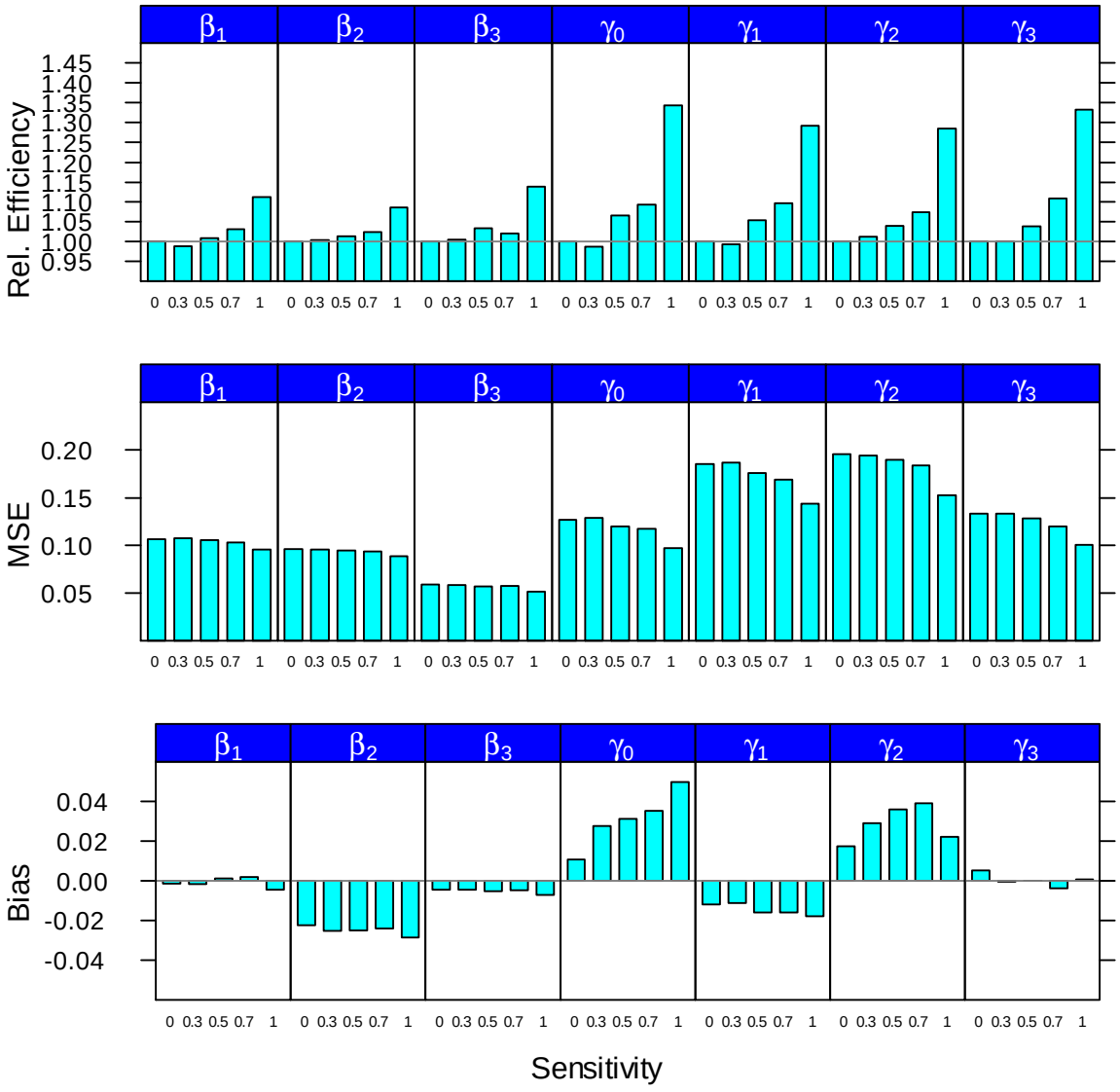


Table 4.6: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1985/-0.0015	0.1983/-0.0017	0.2011/0.0011	0.2019/0.0019	0.1954/-0.0046
	SD		0.3258	0.3276	0.3244	0.3209	0.3089
	MSE		0.1061	0.1073	0.1053	0.1029	0.0955
	Relative Efficiency		**1.0000**	**0.9888**	**1.0081**	**1.0308**	**1.1118**
β_2	Mean/Bias	-0.3	-0.3224/-0.0224	-0.3253/-0.0253	-0.3251/-0.0251	-0.3241/-0.0241	-0.3285/-0.0285
	SD		0.3085	0.3080	0.3064	0.3049	0.2960
	MSE		0.0957	0.0955	0.0945	0.0936	0.0884
	Relative Efficiency		**1.0000**	**1.0032**	**1.0137**	**1.0233**	**1.0860**
β_3	Mean/Bias	0.1	0.0953/-0.0047	0.0954/-0.0046	0.0948/-0.0052	0.0951/-0.0049	0.0929/-0.0071
	SD		0.2421	0.2415	0.2382	0.2397	0.2269
	MSE		0.0586	0.0584	0.0568	0.0575	0.0515
	Relative Efficiency		**1.0000**	**1.0044**	**1.0330**	**1.0202**	**1.1384**
γ_0	Mean/Bias	0.25	0.2608/0.0108	0.2778/0.0278	0.2812/0.0312	0.2853/0.0353	0.2997/0.0497
	SD		0.3560	0.3583	0.3450	0.3407	0.3072
	MSE		0.1269	0.1291	0.1200	0.1173	0.0968
	Relative Efficiency		**1.0000**	**0.9875**	**1.0649**	**1.0923**	**1.3433**
γ_1	Mean/Bias	-0.1	-0.1120/-0.0120	-0.1113/-0.0113	-0.1159/-0.0159	-0.1161/-0.0161	-0.1180/-0.0180
	SD		0.4304	0.4320	0.4192	0.4111	0.3786
	MSE		0.1854	0.1868	0.1760	0.1693	0.1436
	Relative Efficiency		**1.0000**	**0.9925**	**1.0539**	**1.0961**	**1.2925**
γ_2	Mean/Bias	0.5	0.5175/0.0175	0.5292/0.0292	0.5360/0.0360	0.5391/0.0391	0.5222/0.0222
	SD		0.4424	0.4398	0.4339	0.4269	0.3903
	MSE		0.1960	0.1943	0.1896	0.1838	0.1529
	Relative Efficiency		**1.0000**	**1.0117**	**1.0394**	**1.0737**	**1.2844**
γ_3	Mean/Bias	-0.1	-0.0949/0.0051	-0.1005/-0.0005	-0.1000/0.0000	-0.1038/-0.0038	-0.0993/0.0007
	SD		0.3652	0.3651	0.3585	0.3469	0.3164
	MSE		0.1334	0.1333	0.1286	0.1203	0.1001
	Relative Efficiency		**1.0000**	**1.0004**	**1.0376**	**1.1085**	**1.3325**

Figure 4.7: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity



4.3.4 Models using Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$

For models using Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$, the results with 80% specificity ($p_1 = 0.2$) are shown in Appendix Tables A.13 (assuming known sensitivity and specificity) and A.14 (assuming unknown sensitivity and specificity); the results with 90% specificity ($p_1 = 0.1$) are shown in Appendix Tables A.15(assuming known sensitivity and specificity), A.16 (assuming unknown sensitivity and specificity), Appendix Tables A.17(assuming known sensitivity and specificity, 50% subjects with missing diagnostic information), and A.18 (assuming unknown sensitivity and specificity, 50% subjects with missing diagnostic information); and the results with 100% specificity ($p_1 = 0$) are shown in Tables 4.7 (assuming known sensitivity and specificity), 4.8 (assuming unknown sensitivity and specificity), 4.9 (assuming known sensitivity and specificity, 50% subjects with missing diagnostic information), and 4.10 (assuming unknown sensitivity and specificity, 50% subjects with missing diagnostic information). The simulation results using other sets of Weibull distributions are similar. The mean, standard error, and MSE together with the relative efficiency with respect to the model without using the additional information are presented in all tables. Corresponding figures for the relative efficiency, MSE, and bias are presented from Appendix Figures A.13 through A.18, and from Figures 4.8 through 4.11.

Similar observations can be made as for the case for other baseline Weibull distributions. Biases are small when comparing to true values. The relative efficiency can be seen to increase with higher diagnostic sensitivity and specificity. This efficiency gain is higher when the sensitivity and specificity information are known. The extent of such increase, is larger than other baseline Weibull distributions. For example, when both the sensitivity and specificity are known to be 100%, the relative efficiency gain for γ_0 when $h = \frac{1}{3}$ and $k = 4$ is 2.5984, which is much larger than the similar scenarios under other baseline Weibull distributions. If the diagnostic information is missing for some subjects, the efficiency gain is still seen although the extent of such gain is smaller. when both the sensitivity and specificity are known to be 100%, if 50% of subjects have

missing diagnostic information, the relative efficiency gain for γ_0 is 1.7402.

Table 4.7: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1969/-0.0031	0.1930/-0.0070	0.1963/-0.0037	0.1911/-0.0089	0.1955/-0.0045
	SD		0.4068	0.3998	0.3964	0.3847	0.3511
	MSE		0.1655	0.1599	0.1572	0.1481	0.1233
	Relative Efficiency		**1.0000**	**1.0356**	**1.0532**	**1.1182**	**1.3428**
β_2	Mean/Bias	-0.3	-0.3191/-0.0191	-0.3216/-0.0216	-0.3197/-0.0197	-0.3201/-0.0201	-0.3136/-0.0136
	SD		0.3985	0.3876	0.3783	0.3725	0.3403
	MSE		0.1592	0.1507	0.1435	0.1391	0.1160
	Relative Efficiency		**1.0000**	**1.0570**	**1.1101**	**1.1448**	**1.3717**
β_3	Mean/Bias	0.1	0.0953/-0.0047	0.0953/-0.0047	0.0974/-0.0026	0.0981/-0.0019	0.0982/-0.0018
	SD		0.3269	0.3196	0.3117	0.3044	0.2778
	MSE		0.1069	0.1022	0.0971	0.0927	0.0772
	Relative Efficiency		**1.0000**	**1.0462**	**1.1001**	**1.1530**	**1.3848**
γ_0	Mean/Bias	0.25	0.2757/0.0257	0.2630/0.0130	0.2722/0.0222	0.2640/0.0140	0.2630/0.0130
	SD		0.4517	0.4038	0.3723	0.3353	0.2802
	MSE		0.2047	0.1632	0.1391	0.1126	0.0787
	Relative Efficiency		**1.0000**	**1.2515**	**1.4723**	**1.8146**	**2.5984**
γ_1	Mean/Bias	-0.1	-0.1175/-0.0175	-0.0960/0.0040	-0.1073/-0.0073	-0.0965/0.0035	-0.0976/0.0024
	SD		0.5119	0.4820	0.4474	0.4109	0.3500
	MSE		0.2624	0.2324	0.2002	0.1689	0.1225
	Relative Efficiency		**1.0000**	**1.1279**	**1.3091**	**1.5519**	**2.1388**
γ_2	Mean/Bias	0.5	0.5025/0.0025	0.4994/-0.0006	0.4875/-0.0125	0.4976/-0.0024	0.4827/-0.0173
	SD		0.5969	0.5259	0.4781	0.4368	0.3692
	MSE		0.3563	0.2765	0.2287	0.1908	0.1366
	Relative Efficiency		**1.0000**	**1.2885**	**1.5592**	**1.8678**	**2.6135**
γ_3	Mean/Bias	-0.1	-0.1083/-0.0083	-0.1136/-0.0136	-0.1168/-0.0168	-0.1227/-0.0227	-0.1189/-0.0189
	SD		0.4718	0.4193	0.3817	0.3474	0.2946
	MSE		0.2227	0.1760	0.1460	0.1212	0.0872
	Relative Efficiency		**1.0000**	**1.2658**	**1.5275**	**1.8442**	**2.5641**

Figure 4.8: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity

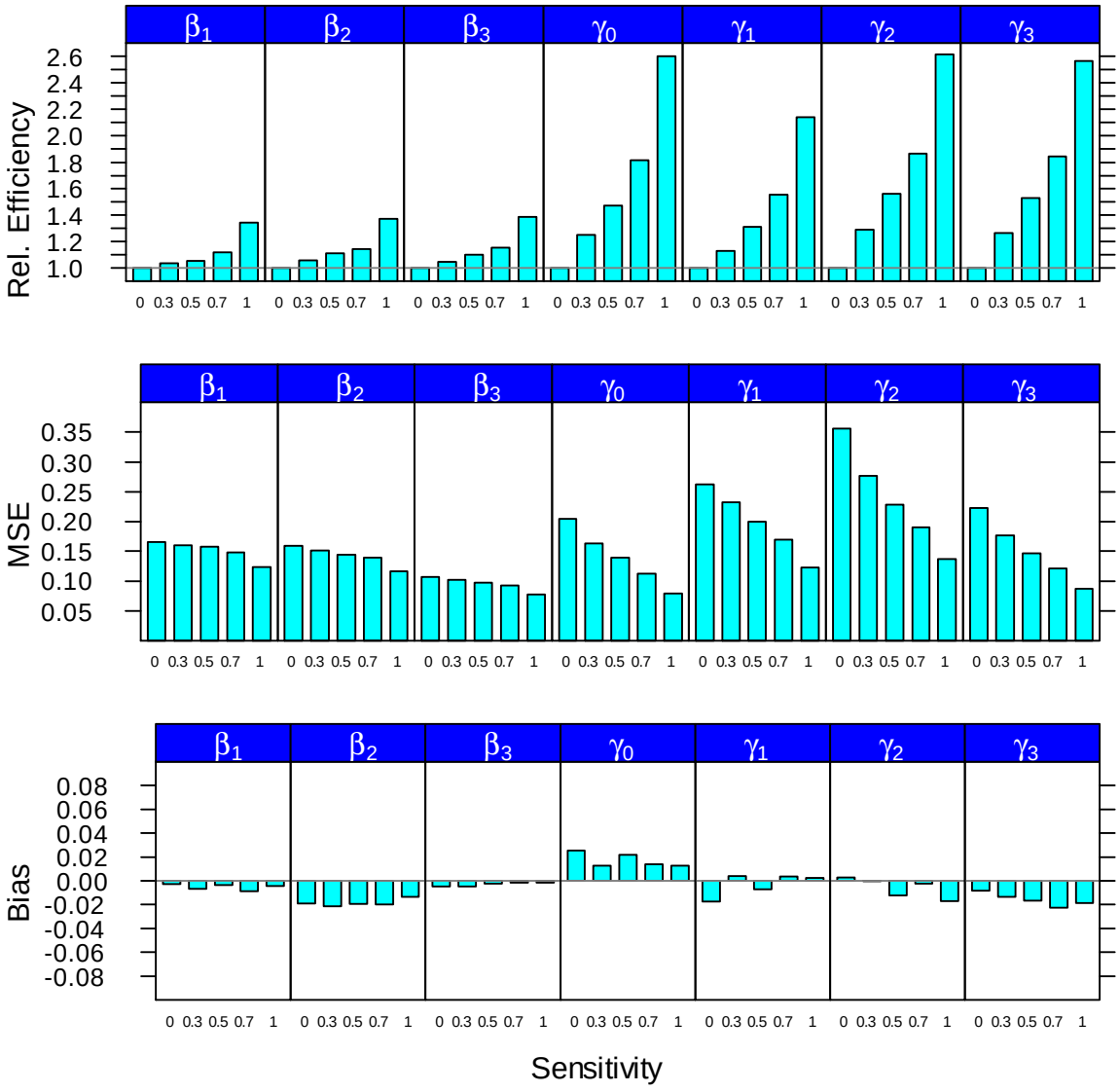


Table 4.8: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1969/-0.0031	0.1959/-0.0041	0.2027/0.0027	0.1963/-0.0037	0.2049/0.0049
	SD		0.4068	0.4069	0.4031	0.3872	0.3582
	MSE		0.1655	0.1656	0.1625	0.1500	0.1284
	Relative Efficiency		**1.0000**	**0.9994**	**1.0186**	**1.1037**	**1.2896**
β_2	Mean/Bias	-0.3	-0.3191/-0.0191	-0.3383/-0.0383	-0.3294/-0.0294	-0.3326/-0.0326	-0.3191/-0.0191
	SD		0.3985	0.3936	0.3852	0.3826	0.3462
	MSE		0.1592	0.1564	0.1492	0.1474	0.1202
	Relative Efficiency		**1.0000**	**1.0251**	**1.0704**	**1.0850**	**1.3254**
β_3	Mean/Bias	0.1	0.0953/-0.0047	0.0993/-0.0007	0.0971/-0.0029	0.1094/0.0094	0.0941/-0.0059
	SD		0.3269	0.3241	0.3124	0.3100	0.2797
	MSE		0.1069	0.1050	0.0976	0.0962	0.0783
	Relative Efficiency		**1.0000**	**1.0173**	**1.0951**	**1.1121**	**1.3658**
γ_0	Mean/Bias	0.25	0.2757/0.0257	0.2896/0.0396	0.3070/0.0570	0.3134/0.0634	0.3406/0.0906
	SD		0.4517	0.4412	0.4244	0.3807	0.3118
	MSE		0.2047	0.1963	0.1834	0.1489	0.1054
	Relative Efficiency		**1.0000**	**1.0480**	**1.1325**	**1.4080**	**2.0984**
γ_1	Mean/Bias	-0.1	-0.1175/-0.0175	-0.1111/-0.0111	-0.1093/-0.0093	-0.1135/-0.0135	-0.1115/-0.0115
	SD		0.5119	0.5117	0.4855	0.4398	0.3737
	MSE		0.2624	0.2620	0.2358	0.1936	0.1398
	Relative Efficiency		**1.0000**	**1.0008**	**1.1119**	**1.3550**	**1.8760**
γ_2	Mean/Bias	0.5	0.5025/0.0025	0.5529/0.0529	0.5593/0.0593	0.5612/0.0612	0.5327/0.0327
	SD		0.5969	0.5980	0.5558	0.5114	0.4125
	MSE		0.3563	0.3604	0.3124	0.2652	0.1712
	Relative Efficiency		**1.0000**	**0.9963**	**1.1535**	**1.3626**	**2.0941**
γ_3	Mean/Bias	-0.1	-0.1083/-0.0083	-0.1159/-0.0159	-0.1298/-0.0298	-0.1315/-0.0315	-0.1323/-0.0323
	SD		0.4718	0.4560	0.4321	0.3781	0.3117
	MSE		0.2227	0.2082	0.1876	0.1440	0.0982
	Relative Efficiency		**1.0000**	**1.0703**	**1.1920**	**1.5568**	**2.2912**

Figure 4.9: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity

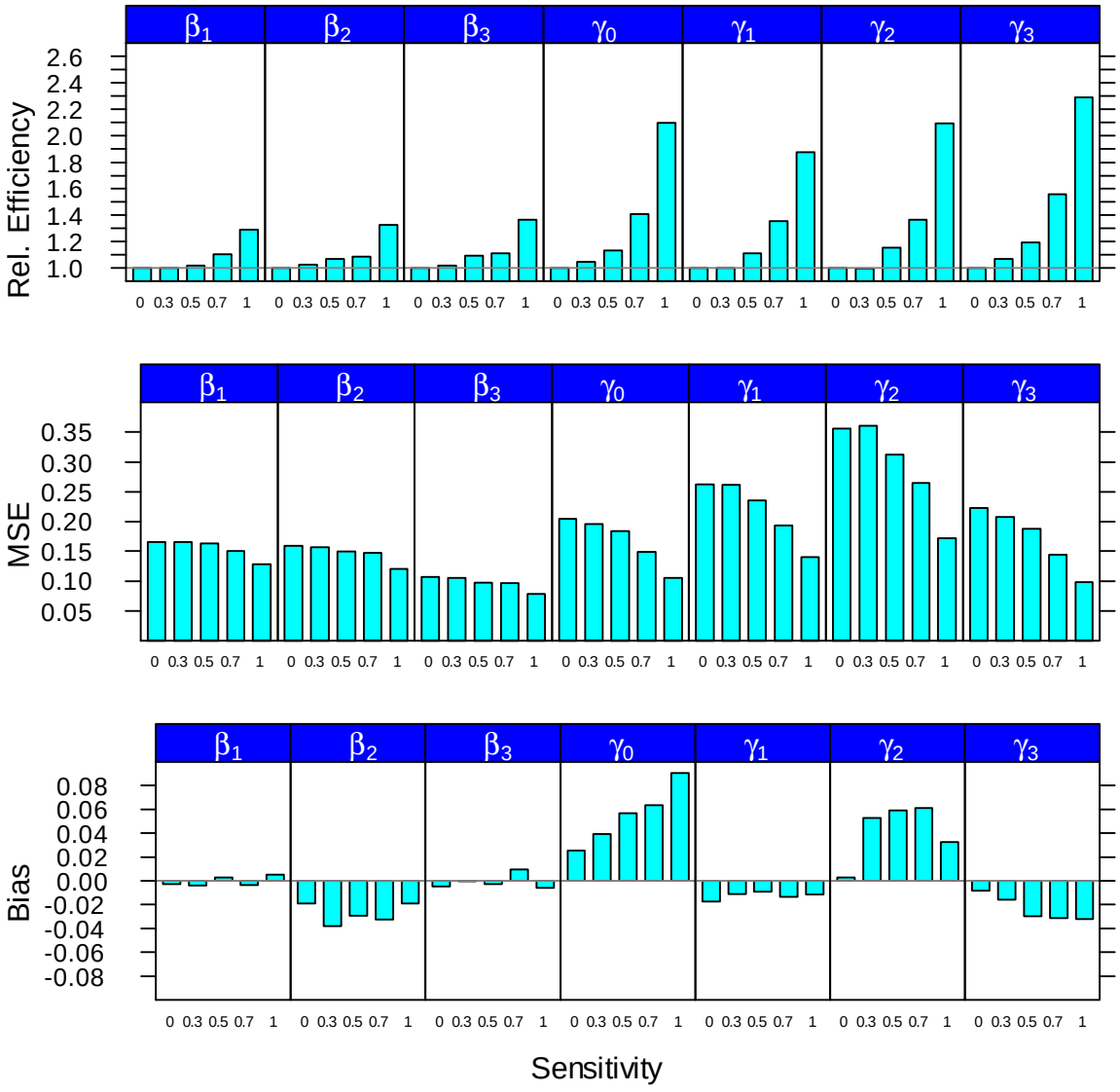


Table 4.9: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1969/-0.0031	0.1964/-0.0036	0.1960/-0.0040	0.1931/-0.0069	0.1923/-0.0077
	SD		0.4068	0.4048	0.4018	0.3965	0.3699
	MSE		0.1655	0.1639	0.1615	0.1573	0.1369
	Relative Efficiency		**1.0000**	**1.0099**	**1.0251**	**1.0526**	**1.2095**
β_2	Mean/Bias	-0.3	-0.3191/-0.0191	-0.3196/-0.0196	-0.3233/-0.0233	-0.3198/-0.0198	-0.3177/-0.0177
	SD		0.3985	0.3940	0.3880	0.3831	0.3594
	MSE		0.1592	0.1556	0.1511	0.1471	0.1295
	Relative Efficiency		**1.0000**	**1.0234**	**1.0550**	**1.0823**	**1.2294**
β_3	Mean/Bias	0.1	0.0953/-0.0047	0.0946/-0.0054	0.0949/-0.0051	0.0975/-0.0025	0.0948/-0.0052
	SD		0.3269	0.3238	0.3198	0.3131	0.2921
	MSE		0.1069	0.1049	0.1023	0.0980	0.0853
	Relative Efficiency		**1.0000**	**1.0189**	**1.0447**	**1.0901**	**1.2527**
γ_0	Mean/Bias	0.25	0.2757/0.0257	0.2702/0.0202	0.2671/0.0171	0.2629/0.0129	0.2592/0.0092
	SD		0.4517	0.4225	0.4064	0.3854	0.3424
	MSE		0.2047	0.1789	0.1655	0.1487	0.1173
	Relative Efficiency		**1.0000**	**1.1430**	**1.2352**	**1.3739**	**1.7402**
γ_1	Mean/Bias	-0.1	-0.1175/-0.0175	-0.1094/-0.0094	-0.1078/-0.0078	-0.1004/-0.0004	-0.1003/-0.0003
	SD		0.5119	0.4976	0.4775	0.4500	0.4140
	MSE		0.2624	0.2477	0.2280	0.2025	0.1714
	Relative Efficiency		**1.0000**	**1.0583**	**1.1495**	**1.2942**	**1.5288**
γ_2	Mean/Bias	0.5	0.5025/0.0025	0.4994/-0.0006	0.5057/0.0057	0.5041/0.0041	0.4965/-0.0035
	SD		0.5969	0.5570	0.5319	0.4884	0.4350
	MSE		0.3563	0.3103	0.2829	0.2385	0.1892
	Relative Efficiency		**1.0000**	**1.1483**	**1.2596**	**1.4938**	**1.8832**
γ_3	Mean/Bias	-0.1	-0.1083/-0.0083	-0.1107/-0.0107	-0.1094/-0.0094	-0.1182/-0.0182	-0.1115/-0.0115
	SD		0.4718	0.4444	0.4222	0.3956	0.3488
	MSE		0.2227	0.1976	0.1784	0.1568	0.1218
	Relative Efficiency		**1.0000**	**1.1271**	**1.2485**	**1.4222**	**1.8293**

Figure 4.10: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information

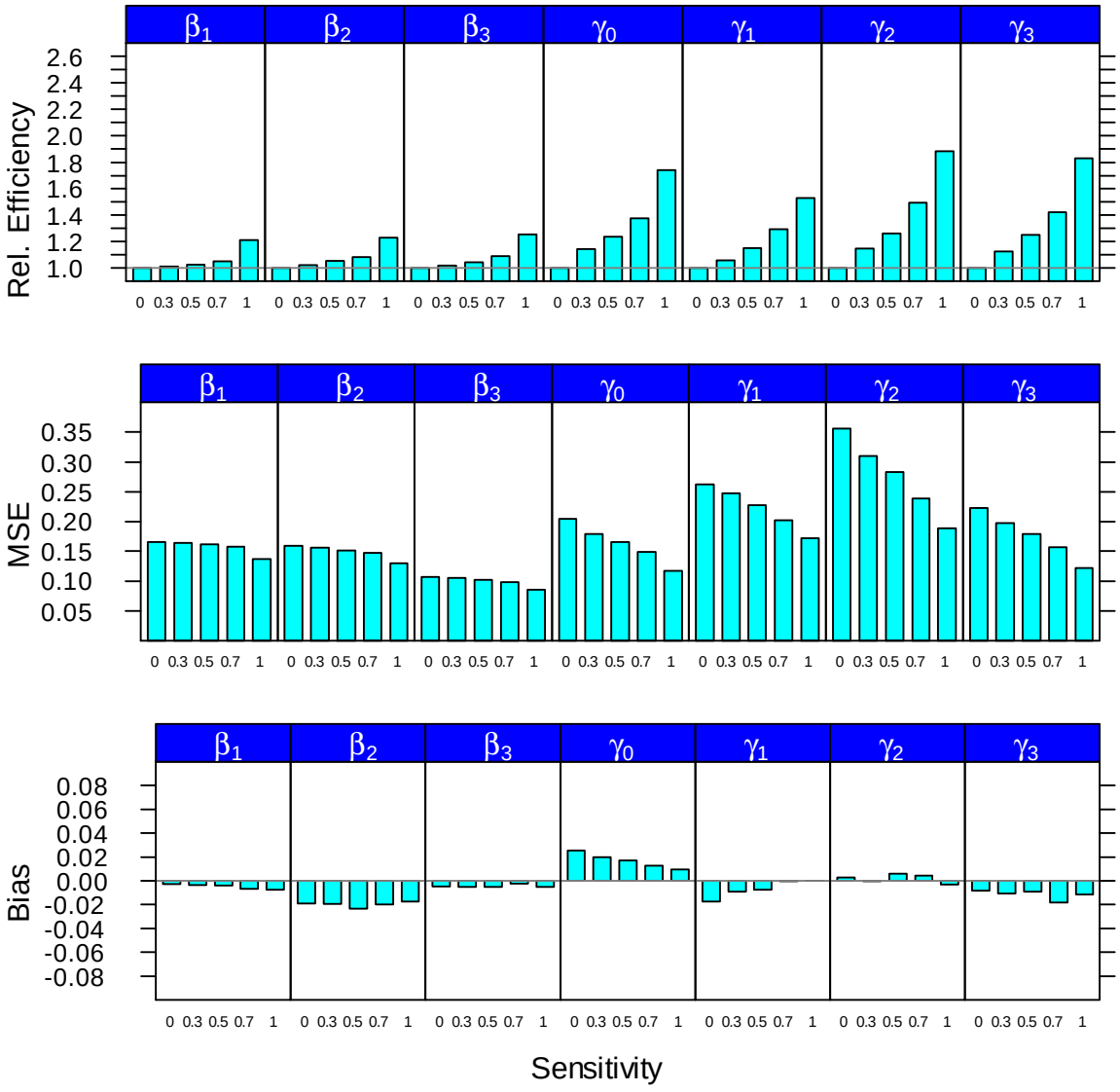
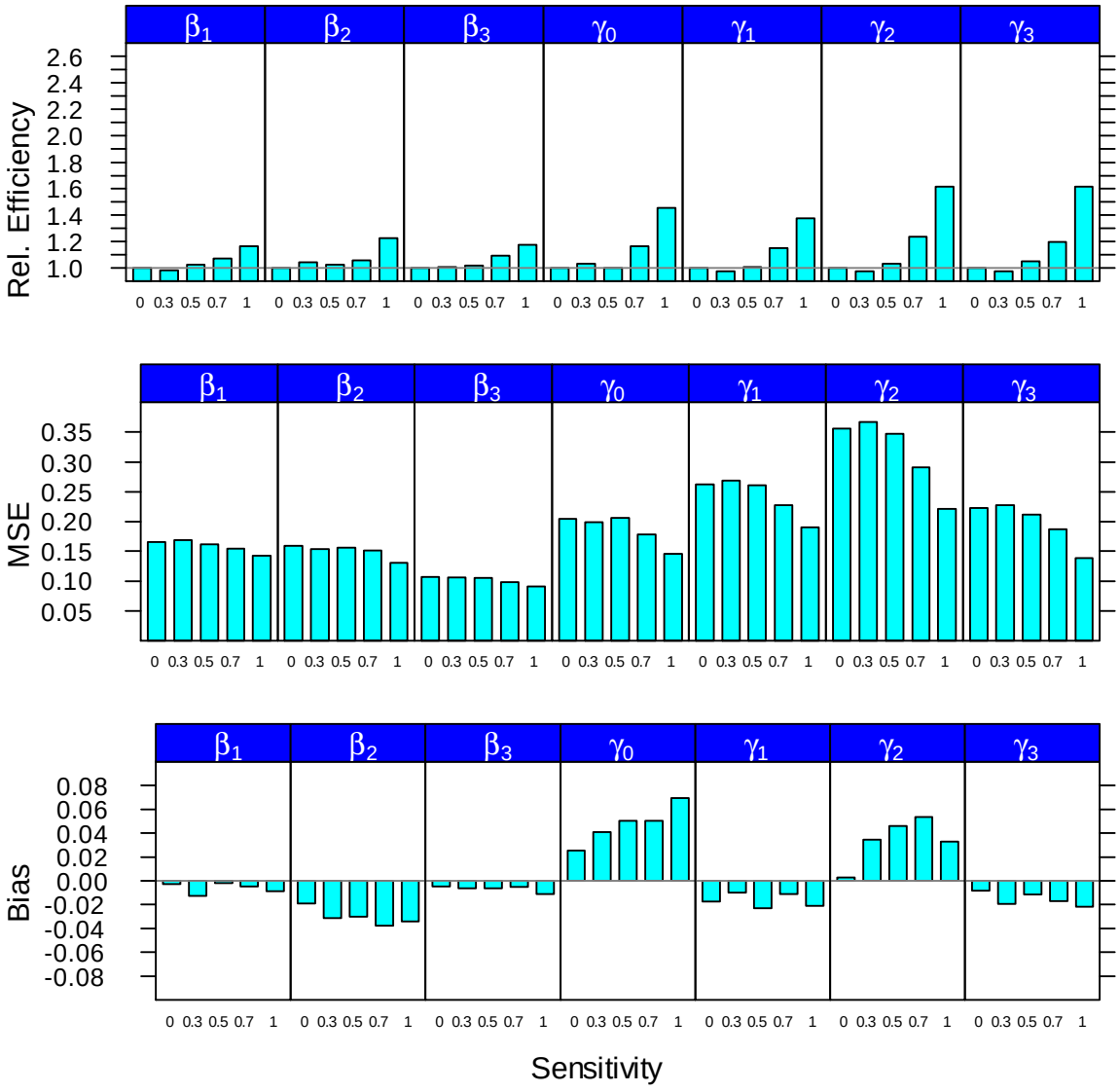


Table 4.10: PH Simulation Results for 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1969/-0.0031	0.1872/-0.0128	0.1977/-0.0023	0.1952/-0.0048	0.1913/-0.0087
	SD		0.4068	0.4102	0.4020	0.3930	0.3771
	MSE		0.1655	0.1684	0.1616	0.1544	0.1423
	Relative Efficiency		**1.0000**	**0.9836**	**1.0243**	**1.0718**	**1.1638**
β_2	Mean/Bias	-0.3	-0.3191/-0.0191	-0.3314/-0.0314	-0.3302/-0.0302	-0.3379/-0.0379	-0.3341/-0.0341
	SD		0.3985	0.3901	0.3938	0.3872	0.3600
	MSE		0.1592	0.1532	0.1560	0.1514	0.1308
	Relative Efficiency		**1.0000**	**1.0435**	**1.0242**	**1.0592**	**1.2254**
β_3	Mean/Bias	0.1	0.0953/-0.0047	0.0936/-0.0064	0.0937/-0.0063	0.0948/-0.0052	0.0886/-0.0114
	SD		0.3269	0.3258	0.3236	0.3128	0.3017
	MSE		0.1069	0.1062	0.1048	0.0979	0.0912
	Relative Efficiency		**1.0000**	**1.0067**	**1.0204**	**1.0920**	**1.1739**
γ_0	Mean/Bias	0.25	0.2757/0.0257	0.2909/0.0409	0.3005/0.0505	0.3003/0.0503	0.3195/0.0695
	SD		0.4517	0.4442	0.4511	0.4188	0.3746
	MSE		0.2047	0.1990	0.2061	0.1780	0.1451
	Relative Efficiency		**1.0000**	**1.0340**	**1.0024**	**1.1630**	**1.4542**
γ_1	Mean/Bias	-0.1	-0.1175/-0.0175	-0.1099/-0.0099	-0.1231/-0.0231	-0.1113/-0.0113	-0.1210/-0.0210
	SD		0.5119	0.5185	0.5099	0.4773	0.4363
	MSE		0.2624	0.2689	0.2606	0.2279	0.1908
	Relative Efficiency		**1.0000**	**0.9748**	**1.0078**	**1.1503**	**1.3766**
γ_2	Mean/Bias	0.5	0.5025/0.0025	0.5348/0.0348	0.5462/0.0462	0.5537/0.0537	0.5332/0.0332
	SD		0.5969	0.6047	0.5876	0.5371	0.4694
	MSE		0.3563	0.3669	0.3475	0.2913	0.2214
	Relative Efficiency		**1.0000**	**0.9745**	**1.0319**	**1.2353**	**1.6175**
γ_3	Mean/Bias	-0.1	-0.1083/-0.0083	-0.1194/-0.0194	-0.1118/-0.0118	-0.1172/-0.0172	-0.1220/-0.0220
	SD		0.4718	0.4772	0.4604	0.4315	0.3709
	MSE		0.2227	0.2281	0.2121	0.1865	0.1380
	Relative Efficiency		**1.0000**	**0.9777**	**1.0503**	**1.1955**	**1.6182**

Figure 4.11: Bias and Relative Efficiency for the PH Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information



4.3.5 Summary results for proposed semi-parametric PH cure rate model

Overall, the point estimators are all close to the true parameters for all the models considered. This indicates the consistency of parameter estimates. On the other hand, in each table, increase of sensitivity p_0 is associated with decrease of SD and MSE, and associated with increase of relative efficiency. Given the same sensitivity level, increase of specificity $1 - p_1$ is associated with decrease of SD, MSE, and thus the increase of relative efficiency. With the same sensitivity and specificity combination, the SD and MSE from models with additional diagnostic information included are consistently less than that from model without additional diagnostic information included.

Greatest relative efficiency is achieved for the hazard function with ($h = \frac{1}{3}$, $k = 4$), followed by ($h = \frac{2}{3}$, $k = 3$), ($h = 1$, $k = 2$), and ($h = 2$, $k = 2$). Comparing this order of greater relative efficiency with the baseline density curves shown in Figure 4.1, a possible explanation is that when event rates are higher earlier in time, there is more event and less censoring, thus less ambiguity in subjects' cured status. Less ambiguity leads to less difference among the three models considered.

When diagnostic procedure results are considered in the model (as shown when $h = \frac{1}{3}$, $k = 4$), the increase in relative efficiency when 50% subjects have missing diagnostic information is less than when 100% subjects have diagnostic results.

The results of the simulations indicate that, with heavy censoring, information from additional diagnostic procedure has the potential to improve the efficiency of the estimator. If we can obtain sensitivity and specificity information of the diagnostic information from external calibration source, this efficiency gain can be potentially higher.

4.4 Evaluation of Semi-Parametric AFT Cure Rate Models

Similar simulation settings were also applied to cure rate models with semiparametric AFT specification. Simulations were run with 100% specificity ($p_1 = 0$), and sensitivity parameters of 0.7 and 1. The same covariates are used in both the "incidence" part and the "latency" part of the model. The latency part is modeled using semi-parametric

AFT model. The covariates for this AFT model are specified as:

$$\log(t_i) = \beta_1 I_{(\text{TRT=1})} + \beta_2 I_{(\text{TRT=2})} + \beta_3 I_{(\text{SEX=Male})} + \epsilon. \tag{4.42}$$

Both latency parameter estimation methods used by Li and Taylor (2002) and Zhang and Peng (2007) were used for the latency parameter estimation. Results from both latency estimation methods will be summarized separately, and also be compared to each other.

4.4.1 Summary of results for the proposed models using Li and Taylor's parameter estimation method

The simulation results using Li and Taylor's method (2002) for parameter estimation are presented in Tables 4.11 to 4.18. The results for models using Weibull distribution with parameter $h = 1$ and $k = 2$ are presented in Tables 4.11 to 4.12. The results for models using Weibull distribution with parameter $h = 2$ and $k = 2$ are presented in Tables 4.13 to 4.14. The results for models using Weibull distribution with parameter $h = \frac{2}{3}$ and $k = 3$ are presented in Tables 4.15 to 4.16. The results for models using Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$ are presented in Tables 4.17 to 4.18. Bias, MSE, and relative efficiency plots for Weibull distribution with parameter $h = 1$ and $k = 2$ are presented in Figures 4.12 and 4.13; for Weibull distribution with parameter $h = 2$ and $k = 2$ are presented in Figures 4.14 and 4.15; for Weibull distribution with parameter $h = \frac{2}{3}$ and $k = 3$ are presented in Figures 4.16 and 4.17. Bias, MSE, and relative efficiency plots for Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$ are not presented because of the scale of bias. The reason of relatively large bias for Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$ will be explored in the discussion below.

To avoid grid search, non-linear minimization as implemented by nlm function in R was used to search for the estimates for (2.20). Same as Li and Taylor, zero-tail completion was used for survival function estimate.

For the hazard functions with parameters ($h = \frac{2}{3}$, $k = 3$), ($h = 1$, $k = 2$), and ($h = 2$, $k = 2$), the point estimators are reasonably close to the true values. The point

estimators for the incidence part appear to be closer to the true values than the latency part. This is likely due to the use of estimating function in (2.20), which could lead to slower convergence to the true values. For the relative efficiency, little or no relative efficiency gain is observed for the hazard function with ($h = 2$, $k = 2$). About 5-30% of relative efficiency gain is observed for the hazard function with ($h = 1$, $k = 2$), with higher relative efficiency gain noted when $p_0 = 1$. About 30-60% relative efficiency gain is observed for the hazard function with ($h = \frac{2}{3}$, $k = 3$), with higher relative efficiency gain noted when $p_0 = 1$. Relative efficiency gain is higher with known sensitivity and specificity parameters than with unknown sensitivity and specificity parameters. Compared to the results from the PH model simulation, the relative efficiency gain is smaller for the AFT simulation in the case of ($h = 2$, $k = 2$), similar in the case of ($h = 1$, $k = 2$), slightly larger in the case of ($h = \frac{2}{3}$, $k = 3$).

Simulation results for the hazard function with ($h = \frac{1}{3}$, $k = 4$) show a different issue. When the sensitivity and specificity parameters are known, large relative efficiency gains are seen. When the sensitivity and specificity parameters are unknown, the relative efficiency gain are seen for most parameters. However, the relative efficiency gain are mainly due to the improvement to the point estimates. When diagnostic information is not included in the model, point estimates are all far from the true value. After including the diagnostic information, the point estimators are improved. Larger improvements are seen in the incidence part than the latency part.

Zhang and Peng (2007) commented on the potential consistency issue of Li and Taylor's method. The consistency could be a result of the grid search for the root of the non-monotonic function $\Psi(\boldsymbol{\beta};g)$. Since the baseline hazard function with ($h = \frac{1}{3}$, $k = 4$) has a high expected censoring rate of 0.4532, another possible explanation to the bias is due to the use of the zero-tail completion when the censoring rate is high. Zero-tail completion is seriously biased when the censoring rate is high. Since equation (2.20) requires the joint estimation of the survival function and $\boldsymbol{\beta}$, the large bias in the survival function estimates also translates to large bias in estimates for $\boldsymbol{\beta}$. Bias in the latency part also leads to bias to the incidence part. Known sensitivity and specificity helps to alleviate the bias, but this issue still remains. This issue is relatively minor (as

seen from the PH simulation results) in the PH model, since $\boldsymbol{\beta}$ is estimated separately from the survival function.

Peng (2003b) discussed about different tail completion scheme under the context of PH models. In the paper, Peng discussed about tail completion by extending the tail using exponential or Weibull distributions. In Table 4.19, we explored the effect of exponential tail completion for the case when hazard function of ($h = \frac{1}{3}$, $k = 4$) is used. Compared Table 4.19 with Table 4.17, the biases in the two tables are showing the same pattern. The effect of tail completion on parameter estimation will be further explored as a part of the future work.

Table 4.11: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2611/0.0611	0.2603/0.0603	0.2593/0.0593
	SD		0.2271	0.2277	0.2215
	MSE		0.0553	0.0555	0.0526
	Relative Efficiency		**1.0000**	**0.9942**	**1.0510**
β_2	Mean/Bias	-0.3	-0.3397/-0.0397	-0.3405/-0.0405	-0.3379/-0.0379
	SD		0.1608	0.1595	0.1551
	MSE		0.0274	0.0271	0.0255
	Relative Efficiency		**1.0000**	**1.0163**	**1.0742**
β_3	Mean/Bias	0.1	0.1201/0.0201	0.1212/0.0212	0.1214/0.0214
	SD		0.1648	0.1632	0.1584
	MSE		0.0276	0.0271	0.0255
	Relative Efficiency		**1.0000**	**1.0198**	**1.0833**
γ_0	Mean/Bias	0.25	0.2663/0.0163	0.2677/0.0177	0.2559/0.0059
	SD		0.3281	0.3142	0.2981
	MSE		0.1079	0.0990	0.0889
	Relative Efficiency		**1.0000**	**1.0904**	**1.2108**
γ_1	Mean/Bias	-0.1	-0.0789/0.0211	-0.0977/0.0023	-0.1045/-0.0045
	SD		0.4013	0.3737	0.3550
	MSE		0.1615	0.1397	0.1261
	Relative Efficiency		**1.0000**	**1.1531**	**1.2775**
γ_2	Mean/Bias	0.5	0.4939/-0.0061	0.4909/-0.0091	0.5038/0.0038
	SD		0.4048	0.3869	0.3645
	MSE		0.1639	0.1497	0.1329
	Relative Efficiency		**1.0000**	**1.0948**	**1.2332**
γ_3	Mean/Bias	-0.1	-0.0995/0.0005	-0.1013/-0.0013	-0.1048/-0.0048
	SD		0.3283	0.3073	0.3011
	MSE		0.1078	0.0944	0.0907
	Relative Efficiency		**1.0000**	**1.1414**	**1.1889**

Figure 4.12: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

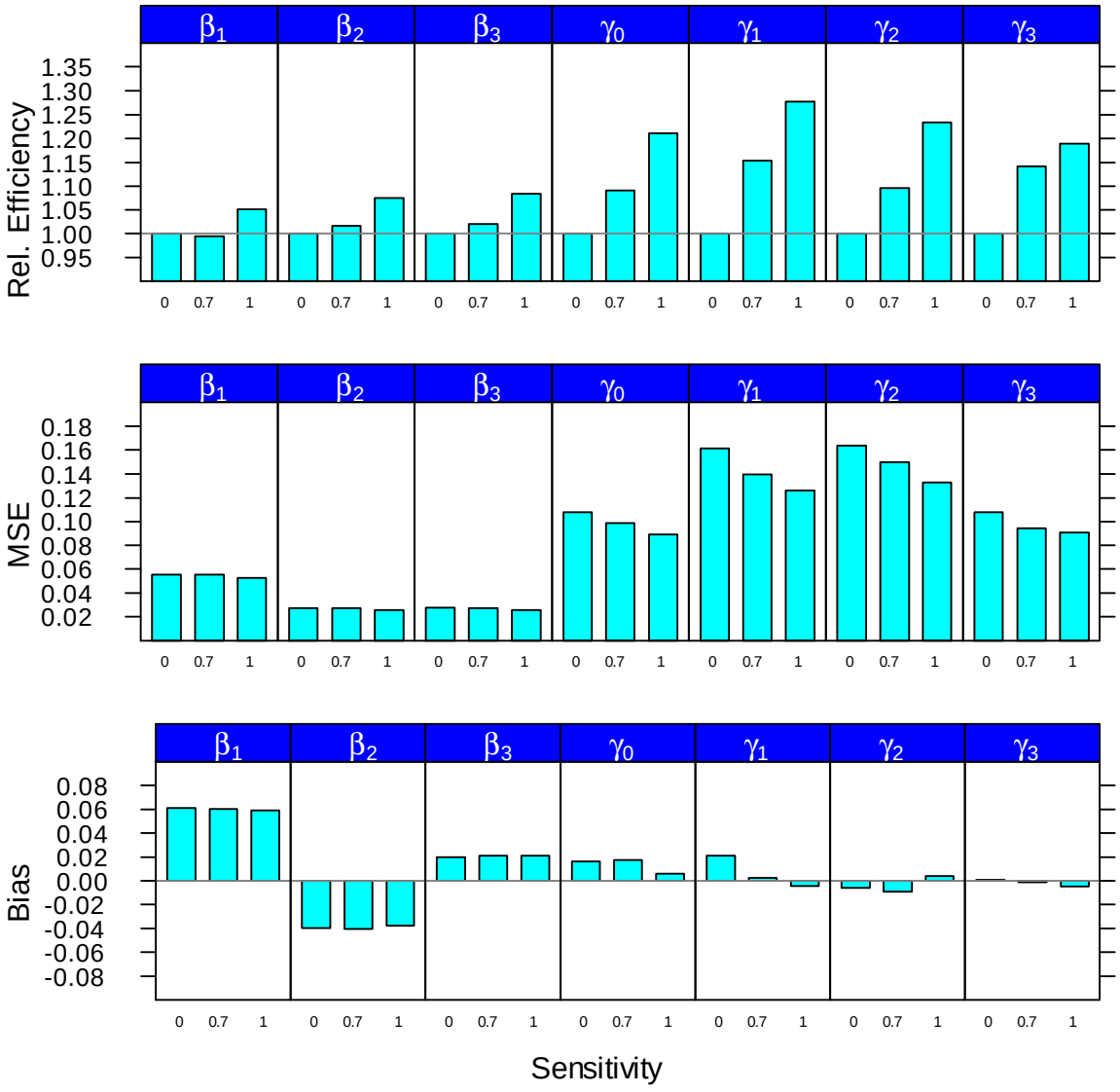


Table 4.12: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2611/0.0611	0.2594/0.0594	0.2588/0.0588
	SD		0.2271	0.2264	0.2194
	MSE		0.0553	0.0548	0.0516
	Relative Efficiency		**1.0000**	**1.0063**	**1.0710**
β_2	Mean/Bias	-0.3	-0.3397/-0.0397	-0.3401/-0.0401	-0.3380/-0.0380
	SD		0.1608	0.1599	0.1552
	MSE		0.0274	0.0272	0.0255
	Relative Efficiency		**1.0000**	**1.0112**	**1.0735**
β_3	Mean/Bias	0.1	0.1201/0.0201	0.1204/0.0204	0.1210/0.0210
	SD		0.1648	0.1628	0.1576
	MSE		0.0276	0.0269	0.0253
	Relative Efficiency		**1.0000**	**1.0254**	**1.0942**
γ_0	Mean/Bias	0.25	0.2663/0.0163	0.2883/0.0383	0.2856/0.0356
	SD		0.3281	0.3219	0.3056
	MSE		0.1079	0.1051	0.0946
	Relative Efficiency		**1.0000**	**1.0389**	**1.1525**
γ_1	Mean/Bias	-0.1	-0.0789/0.0211	-0.0891/0.0109	-0.0964/0.0036
	SD		0.4013	0.3856	0.3639
	MSE		0.1615	0.1488	0.1324
	Relative Efficiency		**1.0000**	**1.0830**	**1.2160**
γ_2	Mean/Bias	0.5	0.4939/-0.0061	0.4926/-0.0074	0.5034/0.0034
	SD		0.4048	0.3953	0.3729
	MSE		0.1639	0.1563	0.1391
	Relative Efficiency		**1.0000**	**1.0483**	**1.1784**
γ_3	Mean/Bias	-0.1	-0.0995/0.0005	-0.1004/-0.0004	-0.1015/-0.0015
	SD		0.3283	0.3147	0.3059
	MSE		0.1078	0.0990	0.0936
	Relative Efficiency		**1.0000**	**1.0882**	**1.1519**

Figure 4.13: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

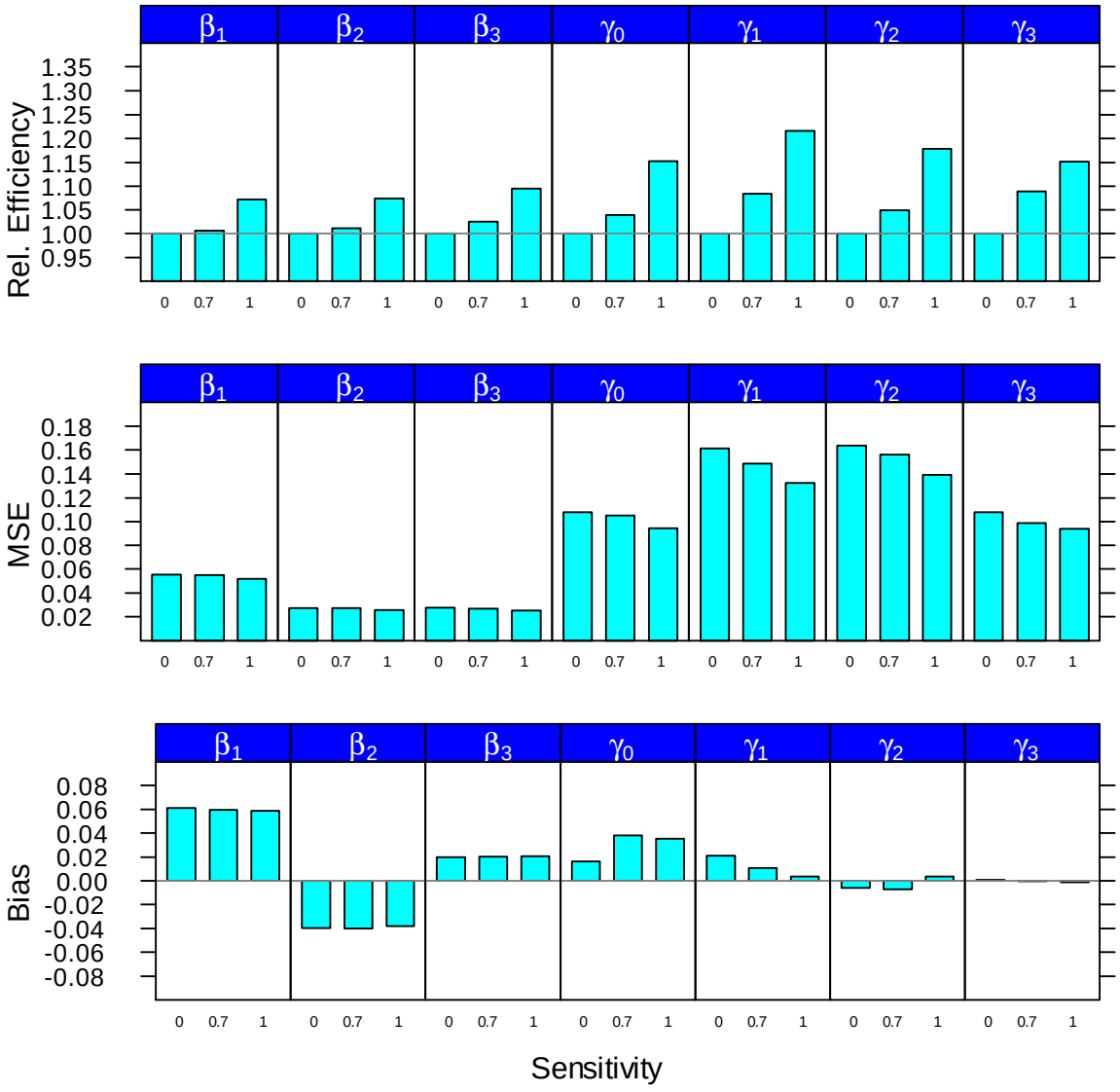


Table 4.13: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2850/0.0850	0.2927/0.0927	0.3011/0.1011
	SD		0.2648	0.2796	0.2927
	MSE		0.0774	0.0868	0.0959
	Relative Efficiency		**1.0000**	**0.8970**	**0.8187**
β_2	Mean/Bias	-0.3	-0.3997/-0.0997	-0.3998/-0.0998	-0.3981/-0.0981
	SD		0.1805	0.1846	0.1868
	MSE		0.0425	0.0440	0.0445
	Relative Efficiency		**1.0000**	**0.9564**	**0.9338**
β_3	Mean/Bias	0.1	0.1468/0.0468	0.1492/0.0492	0.1521/0.0521
	SD		0.1805	0.1872	0.1923
	MSE		0.0348	0.0375	0.0397
	Relative Efficiency		**1.0000**	**0.9294**	**0.8812**
γ_0	Mean/Bias	0.25	0.2536/0.0036	0.2525/0.0025	0.2505/0.0005
	SD		0.3063	0.3010	0.2968
	MSE		0.0938	0.0906	0.0881
	Relative Efficiency		**1.0000**	**1.0355**	**1.0648**
γ_1	Mean/Bias	-0.1	-0.0671/0.0329	-0.0726/0.0274	-0.0904/0.0096
	SD		0.3552	0.3501	0.3531
	MSE		0.1272	0.1233	0.1248
	Relative Efficiency		**1.0000**	**1.0292**	**1.0116**
γ_2	Mean/Bias	0.5	0.4999/-0.0001	0.5046/0.0046	0.5089/0.0089
	SD		0.3681	0.3634	0.3639
	MSE		0.1355	0.1321	0.1325
	Relative Efficiency		**1.0000**	**1.0260**	**1.0227**
γ_3	Mean/Bias	-0.1	-0.0988/0.0012	-0.1012/-0.0012	-0.1050/-0.0050
	SD		0.3193	0.3142	0.3100
	MSE		0.1020	0.0987	0.0961
	Relative Efficiency		**1.0000**	**1.0329**	**1.0610**

Figure 4.14: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

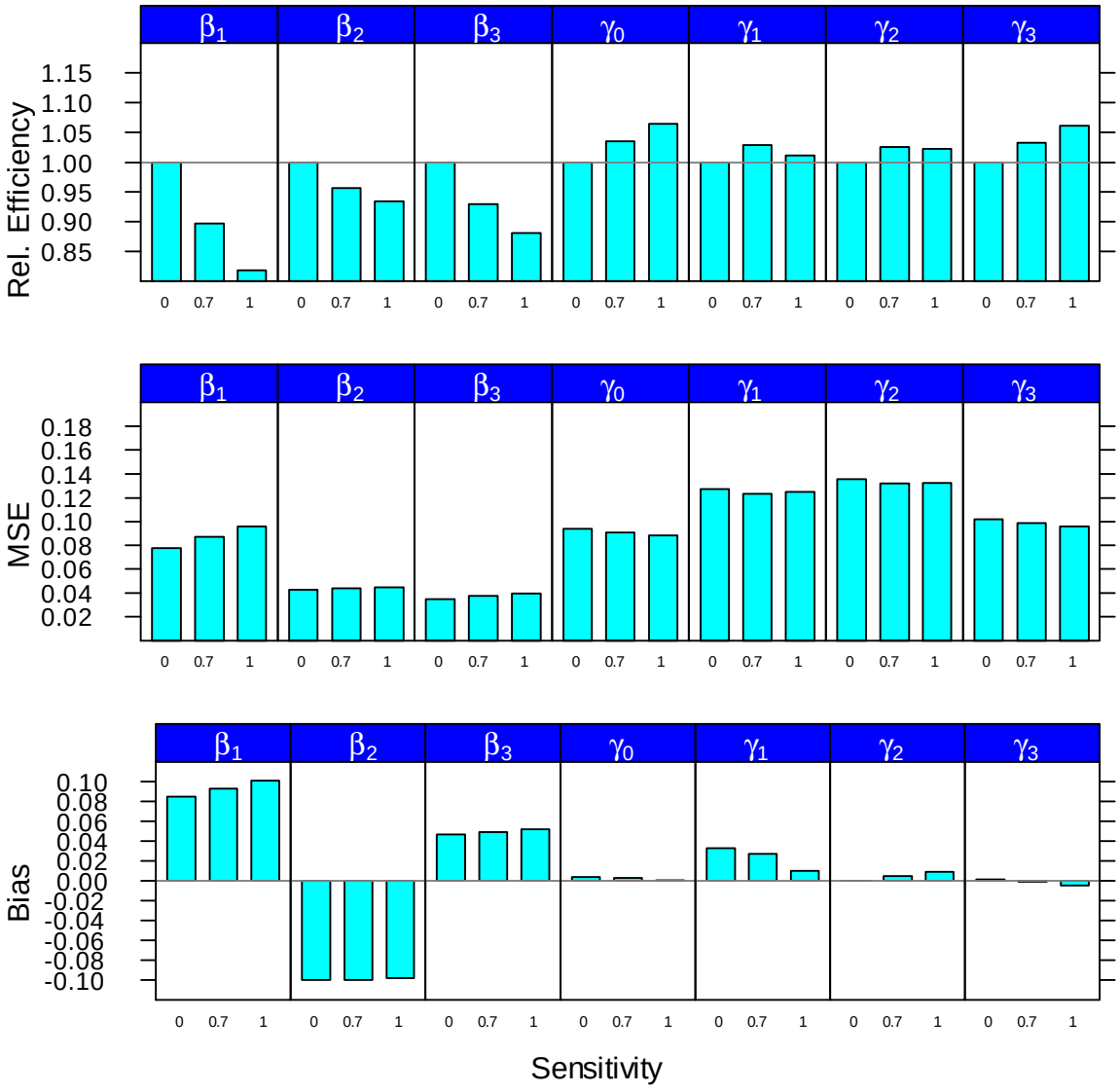


Table 4.14: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2850/0.0850	0.2894/0.0894	0.2971/0.0971
	SD		0.2648	0.2751	0.2852
	MSE		0.0774	0.0837	0.0908
	Relative Efficiency		**1.0000**	**0.9265**	**0.8621**
β_2	Mean/Bias	-0.3	-0.3997/-0.0997	-0.3983/-0.0983	-0.3961/-0.0961
	SD		0.1805	0.1832	0.1852
	MSE		0.0425	0.0432	0.0435
	Relative Efficiency		**1.0000**	**0.9704**	**0.9500**
β_3	Mean/Bias	0.1	0.1468/0.0468	0.1481/0.0481	0.1504/0.0504
	SD		0.1805	0.1836	0.1881
	MSE		0.0348	0.0360	0.0379
	Relative Efficiency		**1.0000**	**0.9670**	**0.9209**
γ_0	Mean/Bias	0.25	0.2536/0.0036	0.2677/0.0177	0.2701/0.0201
	SD		0.3063	0.3025	0.2993
	MSE		0.0938	0.0918	0.0900
	Relative Efficiency		**1.0000**	**1.0252**	**1.0471**
γ_1	Mean/Bias	-0.1	-0.0671/0.0329	-0.0693/0.0307	-0.0840/0.0160
	SD		0.3552	0.3537	0.3535
	MSE		0.1272	0.1260	0.1252
	Relative Efficiency		**1.0000**	**1.0083**	**1.0094**
γ_2	Mean/Bias	0.5	0.4999/-0.0001	0.5014/0.0014	0.5067/0.0067
	SD		0.3681	0.3663	0.3658
	MSE		0.1355	0.1342	0.1338
	Relative Efficiency		**1.0000**	**1.0098**	**1.0125**
γ_3	Mean/Bias	-0.1	-0.0988/0.0012	-0.1009/-0.0009	-0.1029/-0.0029
	SD		0.3193	0.3172	0.3123
	MSE		0.1020	0.1006	0.0975
	Relative Efficiency		**1.0000**	**1.0135**	**1.0456**

Figure 4.15: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

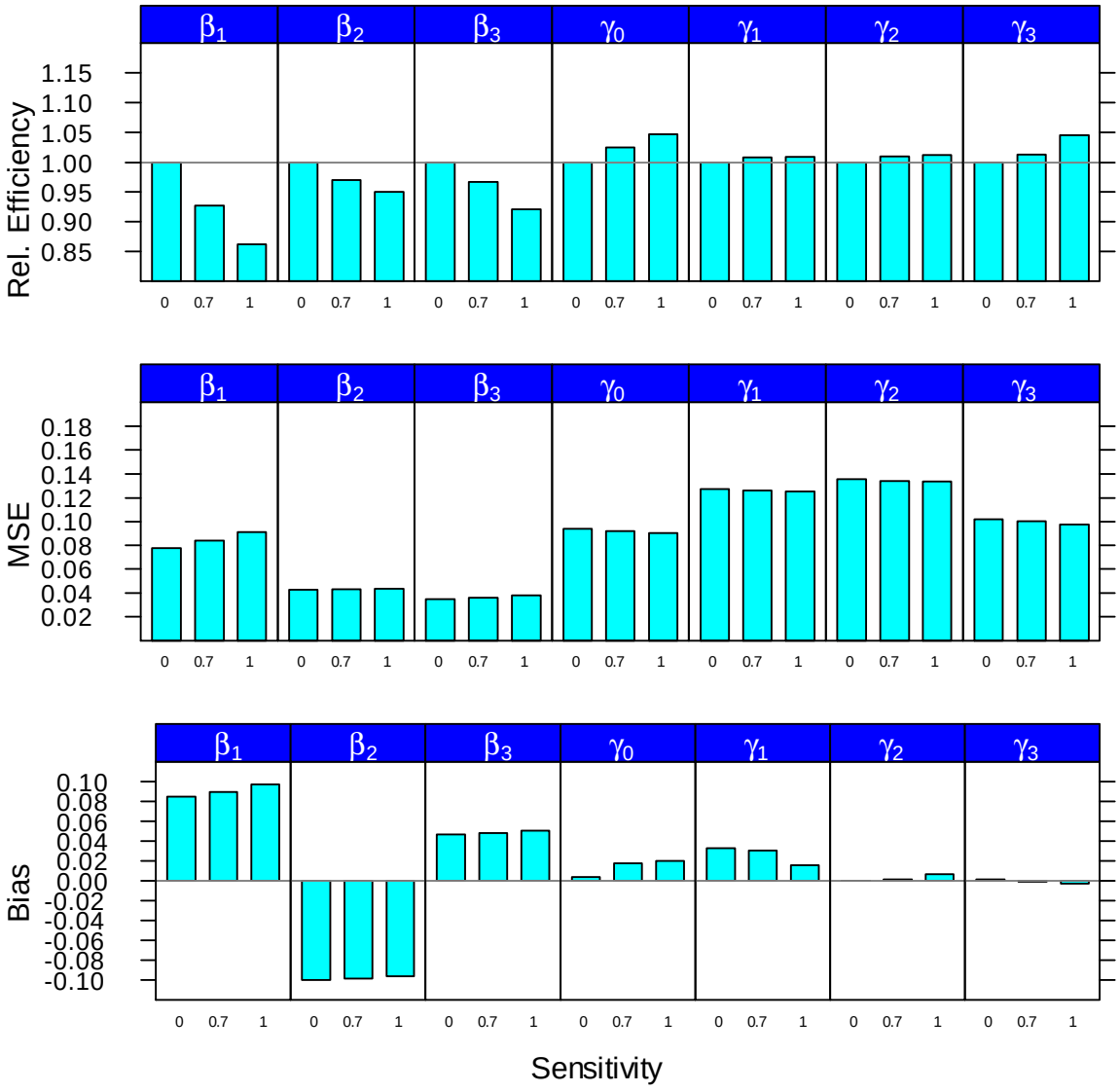


Table 4.15: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.1618/-0.0382	0.1704/-0.0296	0.1776/-0.0224
	SD		0.1642	0.1450	0.1311
	MSE		0.0284	0.0219	0.0177
	Relative Efficiency		**1.0000**	**1.2819**	**1.5684**
β_2	Mean/Bias	-0.3	-0.2556/0.0444	-0.2579/0.0421	-0.2598/0.0402
	SD		0.1071	0.1053	0.1014
	MSE		0.0134	0.0128	0.0119
	Relative Efficiency		**1.0000**	**1.0349**	**1.1159**
β_3	Mean/Bias	0.1	0.0862/-0.0138	0.0881/-0.0119	0.0911/-0.0089
	SD		0.1115	0.1019	0.0937
	MSE		0.0126	0.0105	0.0089
	Relative Efficiency		**1.0000**	**1.1974**	**1.4179**
γ_0	Mean/Bias	0.25	0.2638/0.0138	0.2636/0.0136	0.2603/0.0103
	SD		0.3711	0.3281	0.2937
	MSE		0.1379	0.1079	0.0864
	Relative Efficiency		**1.0000**	**1.2789**	**1.5963**
γ_1	Mean/Bias	-0.1	-0.1398/-0.0398	-0.1225/-0.0225	-0.1085/-0.0085
	SD		0.4816	0.4101	0.3720
	MSE		0.2335	0.1687	0.1385
	Relative Efficiency		**1.0000**	**1.3793**	**1.6758**
γ_2	Mean/Bias	0.5	0.5327/0.0327	0.5172/0.0172	0.5080/0.0080
	SD		0.4470	0.4102	0.3780
	MSE		0.2008	0.1686	0.1430
	Relative Efficiency		**1.0000**	**1.1870**	**1.3981**
γ_3	Mean/Bias	-0.1	-0.1028/-0.0028	-0.1012/-0.0012	-0.0932/0.0068
	SD		0.3888	0.3376	0.3057
	MSE		0.1512	0.1140	0.0935
	Relative Efficiency		**1.0000**	**1.3266**	**1.6181**

Figure 4.16: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

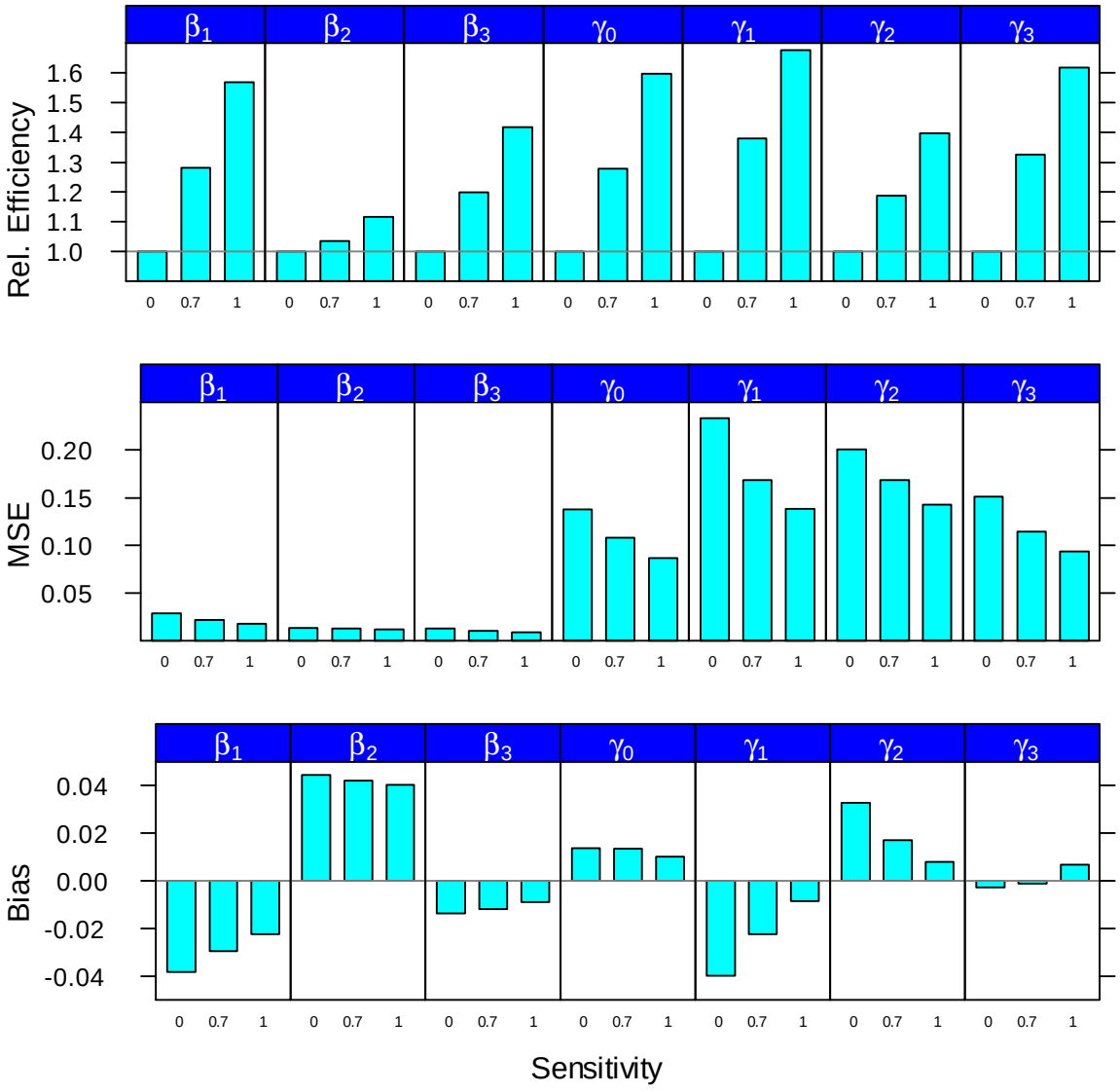


Table 4.16: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.1618/-0.0382	0.1689/-0.0311	0.1778/-0.0222
	SD		0.1642	0.1527	0.1325
	MSE		0.0284	0.0243	0.0180
	Relative Efficiency		**1.0000**	**1.1551**	**1.5361**
β_2	Mean/Bias	-0.3	-0.2556/0.0444	-0.2573/0.0427	-0.2602/0.0398
	SD		0.1071	0.1056	0.1019
	MSE		0.0134	0.0130	0.0120
	Relative Efficiency		**1.0000**	**1.0275**	**1.1032**
β_3	Mean/Bias	0.1	0.0862/-0.0138	0.0875/-0.0125	0.0911/-0.0089
	SD		0.1115	0.1049	0.0943
	MSE		0.0126	0.0112	0.0090
	Relative Efficiency		**1.0000**	**1.1306**	**1.3996**
γ_0	Mean/Bias	0.25	0.2638/0.0138	0.2858/0.0358	0.2921/0.0421
	SD		0.3711	0.3445	0.3068
	MSE		0.1379	0.1200	0.0959
	Relative Efficiency		**1.0000**	**1.1604**	**1.4630**
γ_1	Mean/Bias	-0.1	-0.1398/-0.0398	-0.1230/-0.0230	-0.1051/-0.0051
	SD		0.4816	0.4305	0.3830
	MSE		0.2335	0.1859	0.1467
	Relative Efficiency		**1.0000**	**1.2513**	**1.5812**
γ_2	Mean/Bias	0.5	0.5327/0.0327	0.5257/0.0257	0.5109/0.0109
	SD		0.4470	0.4245	0.3891
	MSE		0.2008	0.1809	0.1515
	Relative Efficiency		**1.0000**	**1.1085**	**1.3194**
γ_3	Mean/Bias	-0.1	-0.1028/-0.0028	-0.1022/-0.0022	-0.0929/0.0071
	SD		0.3888	0.3537	0.3149
	MSE		0.1512	0.1251	0.0992
	Relative Efficiency		**1.0000**	**1.2087**	**1.5250**

Figure 4.17: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

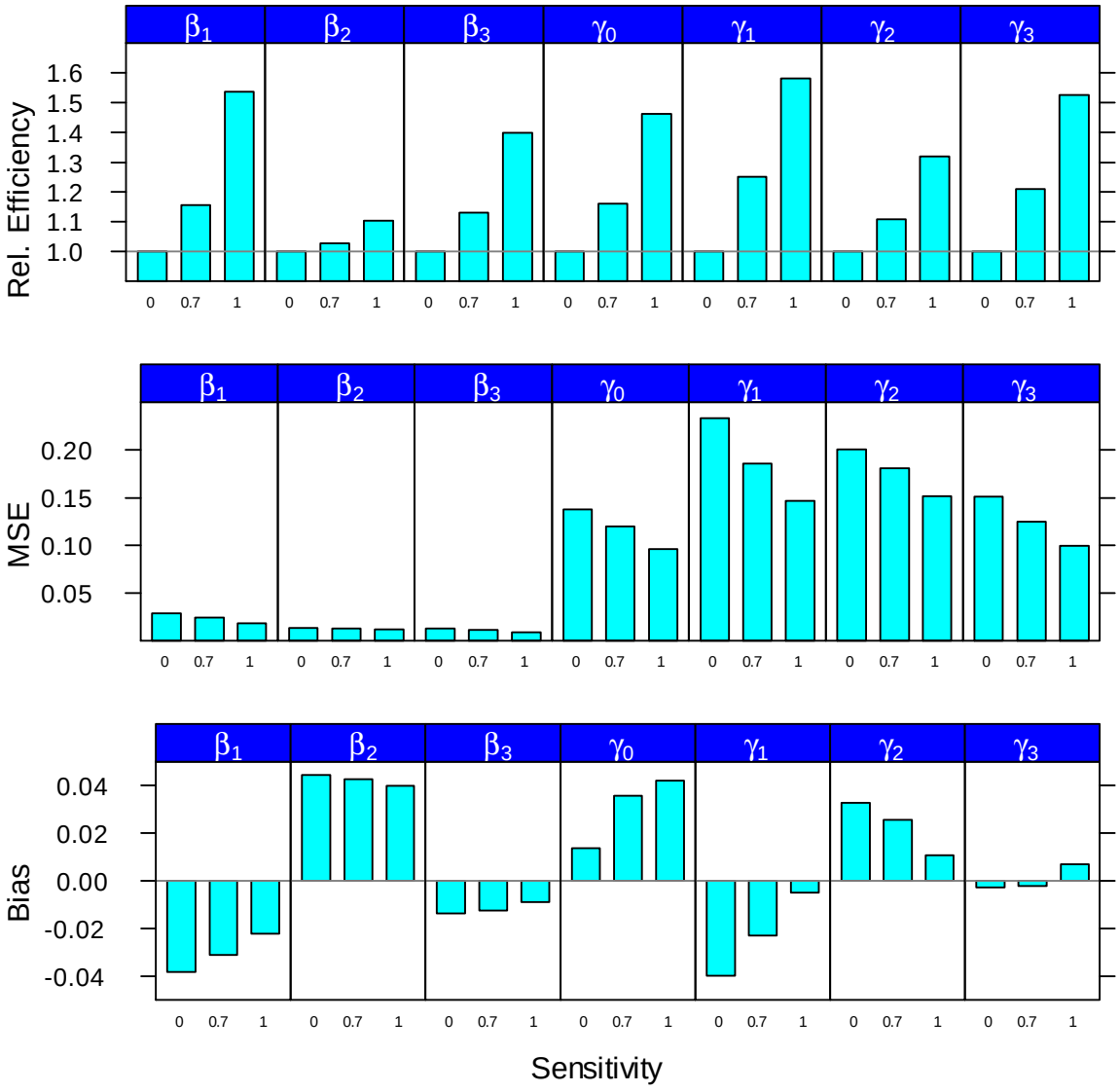


Table 4.17: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	-3.9077/-4.1077	-0.8639/-1.0639	0.0490/-0.1510
	SD		9.6603	3.0770	0.2140
	MSE		110.1936	10.6000	0.0686
	Relative Efficiency		**1.0000**	**9.8562**	**2037.6380**
β_2	Mean/Bias	-0.3	0.1647/0.4647	-0.0481/0.2519	-0.0864/0.2136
	SD		3.3741	0.3791	0.1359
	MSE		11.6004	0.2072	0.0641
	Relative Efficiency		**1.0000**	**79.1941**	**616.4384**
β_3	Mean/Bias	0.1	-1.5819/-1.6819	-0.2316/-0.3316	0.0167/-0.0833
	SD		17.6746	0.9751	0.1453
	MSE		315.2219	1.0608	0.0280
	Relative Efficiency		**1.0000**	**328.5653**	**14805.7533**
γ_0	Mean/Bias	0.25	2.0623/1.8123	0.3358/0.0858	0.2613/0.0113
	SD		5.0876	0.4837	0.2888
	MSE		29.1685	0.2413	0.0836
	Relative Efficiency		**1.0000**	**110.6241**	**310.2485**
γ_1	Mean/Bias	-0.1	-1.4532/-1.3532	-0.4611/-0.3611	-0.1208/-0.0208
	SD		6.0916	0.8188	0.3684
	MSE		38.9388	0.8008	0.1361
	Relative Efficiency		**1.0000**	**55.3546**	**273.4267**
γ_2	Mean/Bias	0.5	1.5880/1.0880	0.6350/0.1350	0.4983/-0.0017
	SD		6.6821	0.5351	0.3731
	MSE		45.8347	0.3045	0.1392
	Relative Efficiency		**1.0000**	**155.9434**	**320.7147**
γ_3	Mean/Bias	-0.1	-0.5948/-0.4948	-0.2347/-0.1347	-0.1306/-0.0306
	SD		3.4610	0.5230	0.3087
	MSE		12.2230	0.2916	0.0962
	Relative Efficiency		**1.0000**	**43.7982**	**125.6771**

Table 4.18: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	-3.9077/-4.1077	-3.0230/-3.2230	-1.9375/-2.1375
	SD		9.6603	5.9145	9.7107
	MSE		110.1936	45.3691	98.8673
	Relative Efficiency		**1.0000**	**2.6677**	**0.9896**
β_2	Mean/Bias	-0.3	0.1647/0.4647	0.2179/0.5179	0.0556/0.3556
	SD		3.3741	1.2169	1.6664
	MSE		11.6004	1.7491	2.9033
	Relative Efficiency		**1.0000**	**7.6881**	**4.0998**
β_3	Mean/Bias	0.1	-1.5819/-1.6819	-0.6437/-0.7437	-0.3134/-0.4134
	SD		17.6746	2.3602	2.7074
	MSE		315.2219	6.1234	7.5010
	Relative Efficiency		**1.0000**	**56.0807**	**42.6176**
γ_0	Mean/Bias	0.25	2.0623/1.8123	2.1580/1.9080	1.4733/1.2233
	SD		5.0876	5.1101	4.3407
	MSE		29.1685	29.7540	20.3380
	Relative Efficiency		**1.0000**	**0.9912**	**1.3738**
γ_1	Mean/Bias	-0.1	-1.4532/-1.3532	-1.6185/-1.5185	-1.2130/-1.1130
	SD		6.0916	5.2963	3.9597
	MSE		38.9388	30.3568	16.9176
	Relative Efficiency		**1.0000**	**1.3229**	**2.3667**
γ_2	Mean/Bias	0.5	1.5880/1.0880	2.5173/2.0173	2.3851/1.8851
	SD		6.6821	6.8653	6.1722
	MSE		45.8347	51.2015	41.6493
	Relative Efficiency		**1.0000**	**0.9474**	**1.1721**
γ_3	Mean/Bias	-0.1	-0.5948/-0.4948	-0.8546/-0.7546	-0.6037/-0.5037
	SD		3.4610	3.5533	3.1267
	MSE		12.2230	13.1951	10.0298
	Relative Efficiency		**1.0000**	**0.9487**	**1.2253**

Table 4.19: Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) - Assume Known Sensitivity and Specificity, Expoential Tail Completion (Using Li and Taylor's Estimation Method)

	Statistics	True Parameter	p_0	
			0.7	1
β_1	Mean/Bias	0.2	-0.8627/-1.0627	0.0489/-0.1511
	SD		3.0475	0.2600
	MSE		10.4169	0.0904
β_2	Mean/Bias	-0.3	-0.0561/0.2439	-0.0782/0.2218
	SD		0.4370	0.1603
	MSE		0.2505	0.0749
β_3	Mean/Bias	0.1	-0.2363/-0.3363	0.0161/-0.0839
	SD		1.1859	0.1716
	MSE		1.5196	0.0365
γ_0	Mean/Bias	0.25	0.3647/0.1147	0.2649/0.0149
	SD		0.4723	0.2871
	MSE		0.2362	0.0827
γ_1	Mean/Bias	-0.1	-0.4738/-0.3738	-0.1180/-0.0180
	SD		0.8095	0.3702
	MSE		0.7950	0.1374
γ_2	Mean/Bias	0.5	0.6295/0.1295	0.4944/-0.0056
	SD		0.5306	0.3728
	MSE		0.2983	0.1390
γ_3	Mean/Bias	-0.1	-0.2428/-0.1428	-0.1304/-0.0304
	SD		0.5198	0.3074
	MSE		0.2905	0.0954

4.4.2 Summary of results for the proposed models using Zhang and Peng's parameter estimation method

The simulation results using Zhang and Peng's method (2007) for parameter estimation are presented in Tables 4.20 to 4.27. The results for models using Weibull distribution with parameter $h = 1$ and $k = 2$ are presented in Tables 4.20 to 4.21. The results for models using Weibull distribution with parameter $h = 2$ and $k = 2$ are presented in Tables 4.22 to 4.23. The results for models using Weibull distribution with parameter $h = \frac{2}{3}$ and $k = 3$ are presented in Tables 4.24 to 4.25. The results for models using Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$ are presented in Tables 4.26 to 4.27. Bias, MSE, and relative efficiency plots for Weibull distribution with parameter $h = 1$ and $k = 2$ are presented in Figures 4.18 and 4.19; for Weibull distribution with parameter $h = 2$ and $k = 2$ are presented in Figures 4.20 and 4.21; for Weibull distribution with parameter $h = \frac{2}{3}$ and $k = 3$ are presented in Figures 4.22 and 4.23; for Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$ are presented in Figures 4.24 and 4.25.

Instead of linear programming as suggested by Zhang and Peng (2007), non-linear minimization as implemented by nlm function in R was used to search for the estimates for (2.26). Zero-tail completion is used for survival function estimate.

For all hazard functions, the point estimators are close to the true values. For the relative efficiency, about 5-15% efficiency gain is observed for the hazard function with ($h = 2$, $k = 2$), about 5-30% of relative efficiency gain is observed for the hazard function with ($h = 1$, $k = 2$), about 15-50% relative efficiency gain is observed for the hazard function with ($h = \frac{2}{3}$, $k = 3$), and >50% relative efficiency gain for the hazard function with ($h = \frac{1}{3}$, $k = 4$). Higher relative efficiency gains are noted when $p_0 = 1$, and this gain is larger with known sensitivity and specificity parameters than with unknown sensitivity and specificity parameters.

Table 4.20: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2083/0.0083	0.2075/0.0075	0.2104/0.0104
	SD		0.1711	0.1671	0.1590
	MSE		0.0293	0.0280	0.0254
	Relative Efficiency		**1.0000**	**1.0474**	**1.1568**
β_2	Mean/Bias	-0.3	-0.2947/0.0053	-0.2949/0.0051	-0.2945/0.0055
	SD		0.1523	0.1518	0.1443
	MSE		0.0232	0.0231	0.0208
	Relative Efficiency		**1.0000**	**1.0059**	**1.1142**
β_3	Mean/Bias	0.1	0.0978/-0.0022	0.0983/-0.0017	0.0985/-0.0015
	SD		0.1312	0.1283	0.1236
	MSE		0.0172	0.0165	0.0153
	Relative Efficiency		**1.0000**	**1.0469**	**1.1280**
γ_0	Mean/Bias	0.25	0.2643/0.0143	0.2654/0.0154	0.2577/0.0077
	SD		0.3337	0.3169	0.2982
	MSE		0.1115	0.1007	0.0890
	Relative Efficiency		**1.0000**	**1.1086**	**1.2521**
γ_1	Mean/Bias	-0.1	-0.1119/-0.0119	-0.1100/-0.0100	-0.1044/-0.0044
	SD		0.4170	0.3837	0.3562
	MSE		0.1740	0.1474	0.1269
	Relative Efficiency		**1.0000**	**1.1806**	**1.3700**
γ_2	Mean/Bias	0.5	0.5012/0.0012	0.4983/-0.0017	0.5044/0.0044
	SD		0.4110	0.3877	0.3654
	MSE		0.1689	0.1503	0.1335
	Relative Efficiency		**1.0000**	**1.1238**	**1.2648**
γ_3	Mean/Bias	-0.1	-0.1035/-0.0035	-0.1069/-0.0069	-0.1073/-0.0073
	SD		0.3394	0.3084	0.2984
	MSE		0.1152	0.0951	0.0891
	Relative Efficiency		**1.0000**	**1.2112**	**1.2930**

Figure 4.18: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

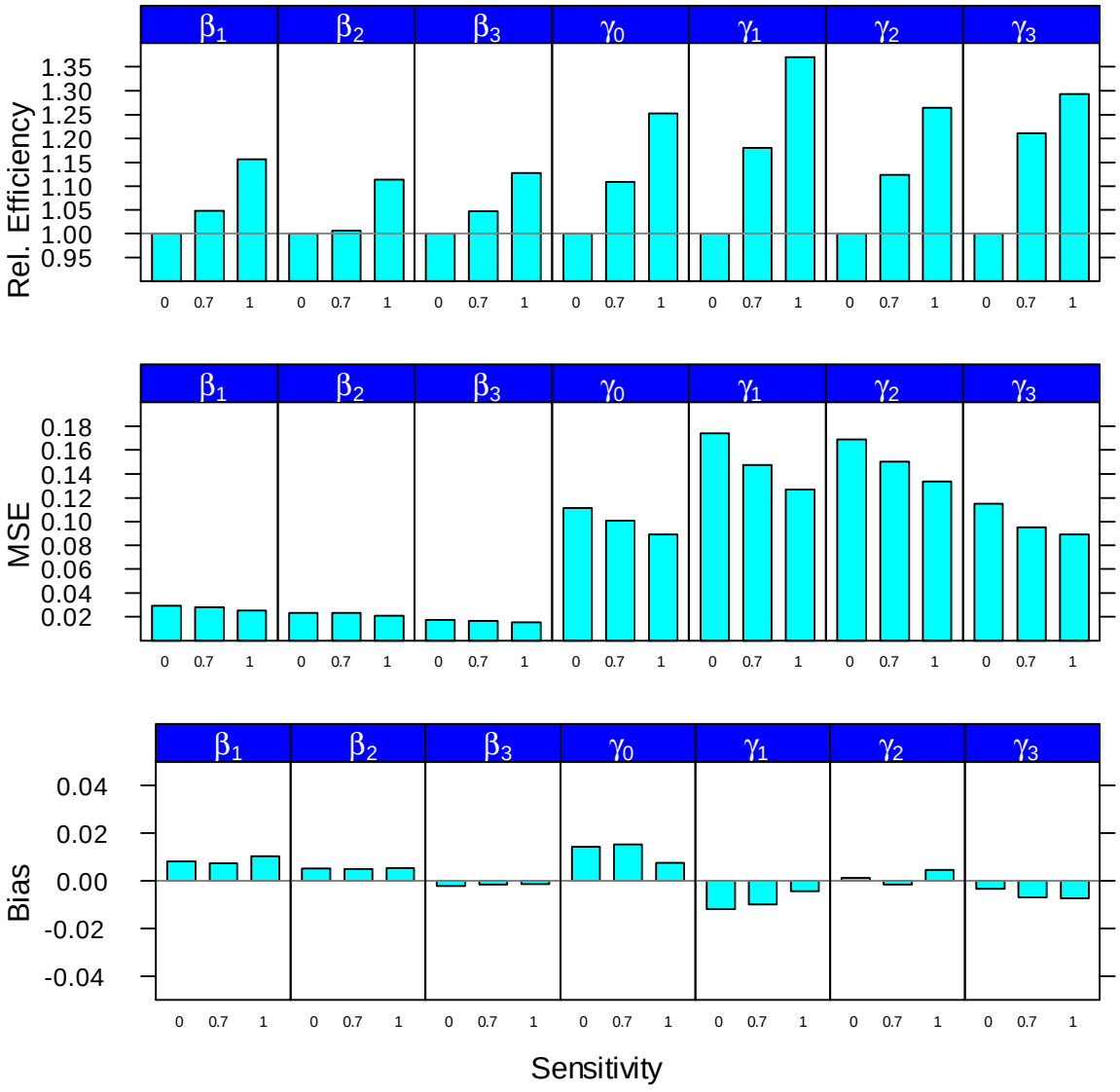


Table 4.21: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2083/0.0083	0.2076/0.0076	0.2115/0.0115
	SD		0.1711	0.1703	0.1615
	MSE		0.0293	0.0290	0.0262
	Relative Efficiency		**1.0000**	**1.0093**	**1.1223**
β_2	Mean/Bias	-0.3	-0.2947/0.0053	-0.2962/0.0038	-0.2954/0.0046
	SD		0.1523	0.1533	0.1454
	MSE		0.0232	0.0235	0.0212
	Relative Efficiency		**1.0000**	**0.9864**	**1.0968**
β_3	Mean/Bias	0.1	0.0978/-0.0022	0.0974/-0.0026	0.0986/-0.0014
	SD		0.1312	0.1294	0.1245
	MSE		0.0172	0.0167	0.0155
	Relative Efficiency		**1.0000**	**1.0291**	**1.1114**
γ_0	Mean/Bias	0.25	0.2643/0.0143	0.2853/0.0353	0.2829/0.0329
	SD		0.3337	0.3245	0.3037
	MSE		0.1115	0.1065	0.0933
	Relative Efficiency		**1.0000**	**1.0572**	**1.2073**
γ_1	Mean/Bias	-0.1	-0.1119/-0.0119	-0.1091/-0.0091	-0.1011/-0.0011
	SD		0.4170	0.3958	0.3650
	MSE		0.1740	0.1567	0.1333
	Relative Efficiency		**1.0000**	**1.1098**	**1.3047**
γ_2	Mean/Bias	0.5	0.5012/0.0012	0.4973/-0.0027	0.5041/0.0041
	SD		0.4110	0.3998	0.3739
	MSE		0.1689	0.1599	0.1398
	Relative Efficiency		**1.0000**	**1.0565**	**1.2083**
γ_3	Mean/Bias	-0.1	-0.1035/-0.0035	-0.1053/-0.0053	-0.1036/-0.0036
	SD		0.3394	0.3220	0.3035
	MSE		0.1152	0.1037	0.0922
	Relative Efficiency		**1.0000**	**1.1110**	**1.2499**

Figure 4.19: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

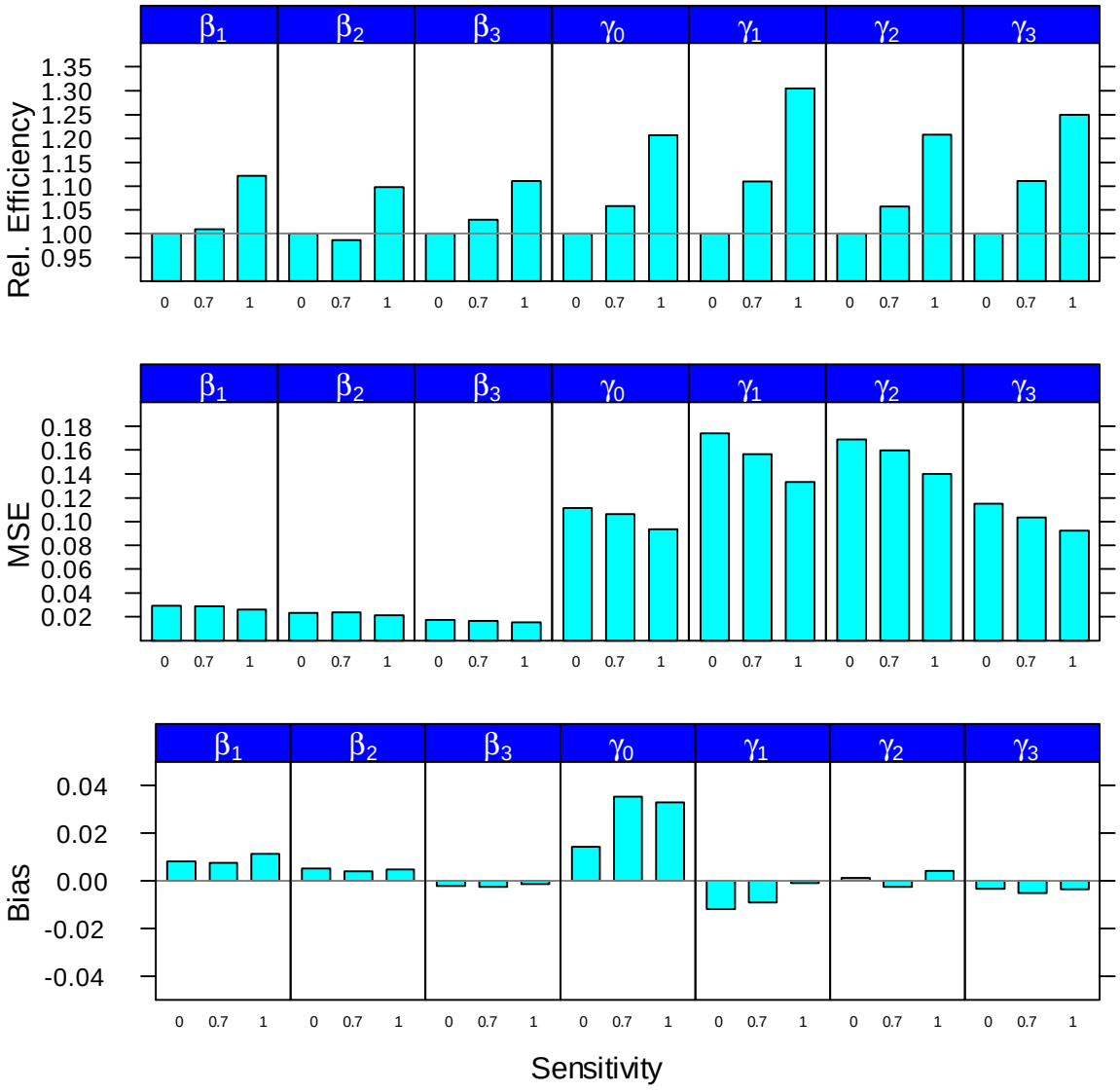


Table 4.22: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2003/0.0003	0.2016/0.0016	0.1998/-0.0002
	SD		0.1548	0.1531	0.1490
	MSE		0.0240	0.0234	0.0222
	Relative Efficiency		**1.0000**	**1.0222**	**1.0789**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3035/-0.0035	-0.3037/-0.0037
	SD		0.1432	0.1421	0.1392
	MSE		0.0205	0.0202	0.0194
	Relative Efficiency		**1.0000**	**1.0146**	**1.0573**
β_3	Mean/Bias	0.1	0.1012/0.0012	0.1020/0.0020	0.1021/0.0021
	SD		0.1191	0.1174	0.1162
	MSE		0.0142	0.0138	0.0135
	Relative Efficiency		**1.0000**	**1.0306**	**1.0510**
γ_0	Mean/Bias	0.25	0.2492/-0.0008	0.2497/-0.0003	0.2519/0.0019
	SD		0.3130	0.3055	0.2979
	MSE		0.0980	0.0933	0.0888
	Relative Efficiency		**1.0000**	**1.0497**	**1.1036**
γ_1	Mean/Bias	-0.1	-0.0816/0.0184	-0.0844/0.0156	-0.0905/0.0095
	SD		0.3772	0.3625	0.3549
	MSE		0.1426	0.1316	0.1261
	Relative Efficiency		**1.0000**	**1.0827**	**1.1294**
γ_2	Mean/Bias	0.5	0.5154/0.0154	0.5126/0.0126	0.5088/0.0088
	SD		0.3770	0.3679	0.3641
	MSE		0.1424	0.1355	0.1326
	Relative Efficiency		**1.0000**	**1.0497**	**1.0721**
γ_3	Mean/Bias	-0.1	-0.1080/-0.0080	-0.1084/-0.0084	-0.1060/-0.0060
	SD		0.3284	0.3218	0.3113
	MSE		0.1079	0.1036	0.0969
	Relative Efficiency		**1.0000**	**1.0416**	**1.1134**

Figure 4.20: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

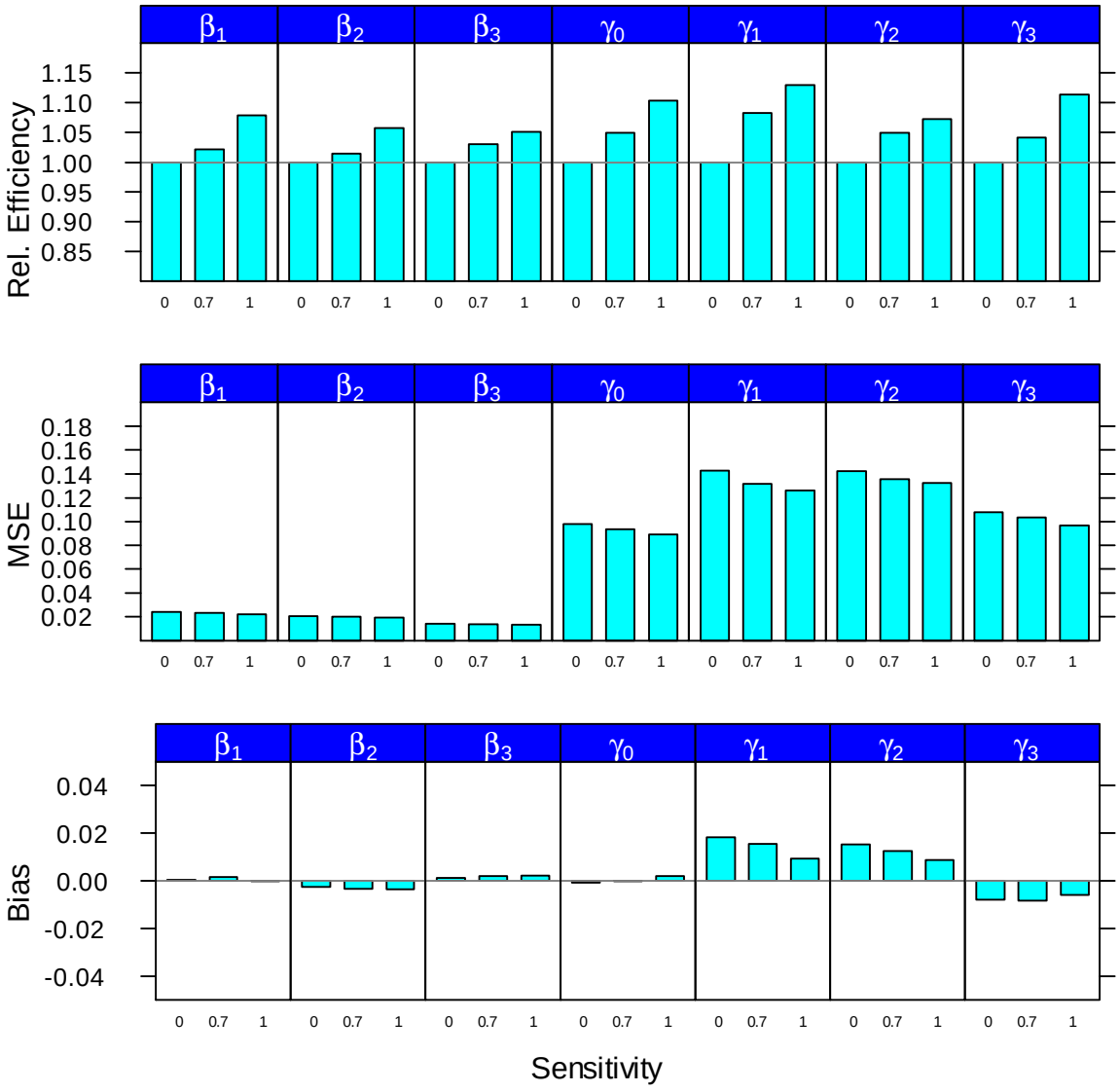


Table 4.23: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2003/0.0003	0.2010/0.0010	0.2007/0.0007
	SD		0.1548	0.1543	0.1484
	MSE		0.0240	0.0238	0.0220
	Relative Efficiency		**1.0000**	**1.0069**	**1.0873**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3053/-0.0053	-0.3048/-0.0048
	SD		0.1432	0.1410	0.1394
	MSE		0.0205	0.0199	0.0194
	Relative Efficiency		**1.0000**	**1.0305**	**1.0551**
β_3	Mean/Bias	0.1	0.1012/0.0012	0.1035/0.0035	0.1028/0.0028
	SD		0.1191	0.1179	0.1170
	MSE		0.0142	0.0139	0.0137
	Relative Efficiency		**1.0000**	**1.0213**	**1.0372**
γ_0	Mean/Bias	0.25	0.2492/-0.0008	0.2630/0.0130	0.2627/0.0127
	SD		0.3130	0.3089	0.3014
	MSE		0.0980	0.0956	0.0910
	Relative Efficiency		**1.0000**	**1.0265**	**1.0783**
γ_1	Mean/Bias	-0.1	-0.0816/0.0184	-0.0850/0.0150	-0.0876/0.0124
	SD		0.3772	0.3705	0.3595
	MSE		0.1426	0.1375	0.1294
	Relative Efficiency		**1.0000**	**1.0367**	**1.1007**
γ_2	Mean/Bias	0.5	0.5154/0.0154	0.5128/0.0128	0.5112/0.0112
	SD		0.3770	0.3723	0.3664
	MSE		0.1424	0.1388	0.1344
	Relative Efficiency		**1.0000**	**1.0252**	**1.0588**
γ_3	Mean/Bias	-0.1	-0.1080/-0.0080	-0.1095/-0.0095	-0.1007/-0.0007
	SD		0.3284	0.3220	0.3142
	MSE		0.1079	0.1038	0.0987
	Relative Efficiency		**1.0000**	**1.0407**	**1.0929**

Figure 4.21: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

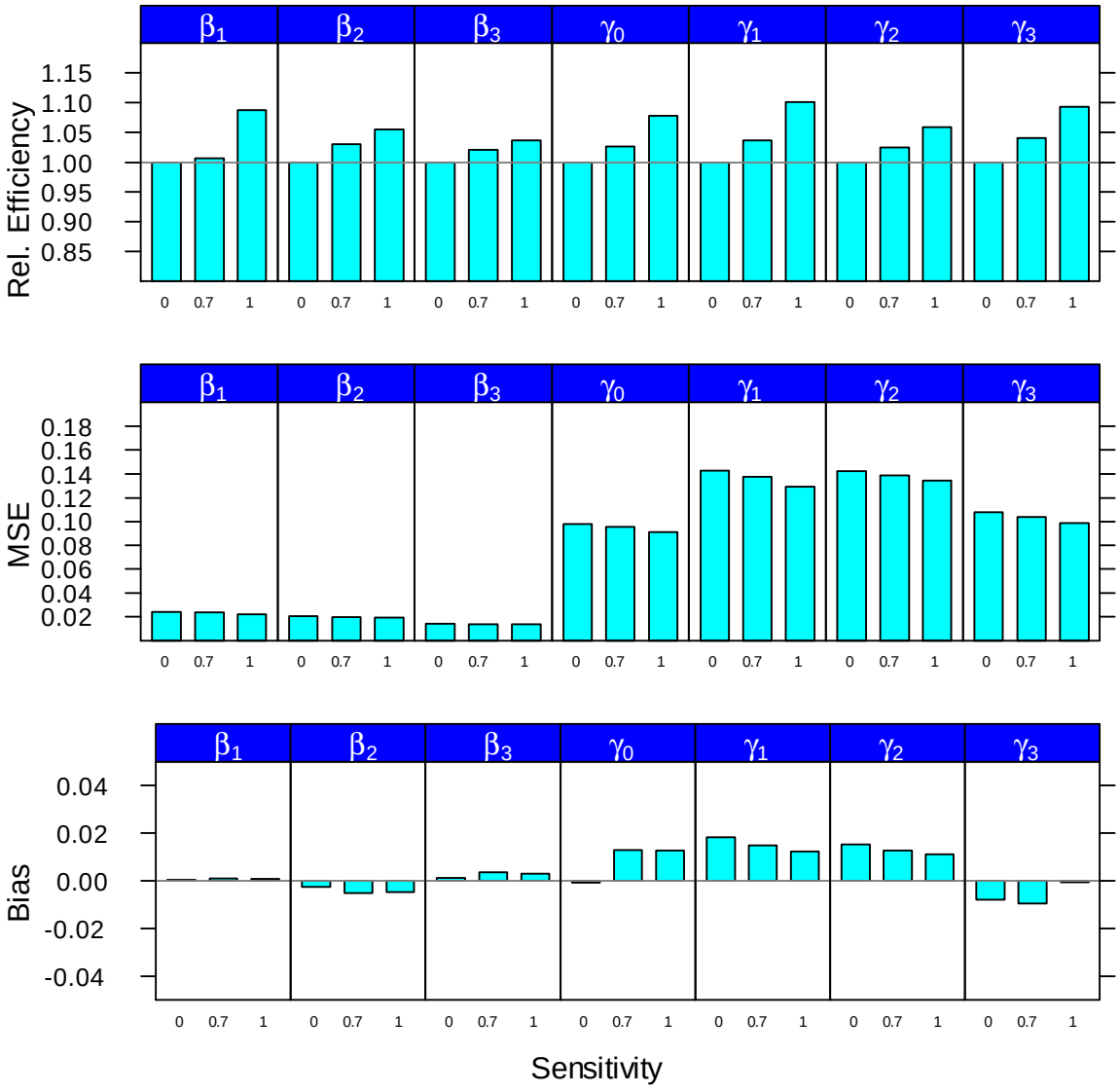


Table 4.24: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2004/0.0004	0.2019/0.0019	0.2031/0.0031
	SD		0.1245	0.1213	0.1147
	MSE		0.0155	0.0147	0.0132
	Relative Efficiency		**1.0000**	**1.0542**	**1.1791**
β_2	Mean/Bias	-0.3	-0.2942/0.0058	-0.2965/0.0035	-0.2954/0.0046
	SD		0.1085	0.1080	0.1038
	MSE		0.0118	0.0117	0.0108
	Relative Efficiency		**1.0000**	**1.0087**	**1.0909**
β_3	Mean/Bias	0.1	0.1018/0.0018	0.1011/0.0011	0.1023/0.0023
	SD		0.0925	0.0902	0.0855
	MSE		0.0086	0.0081	0.0073
	Relative Efficiency		**1.0000**	**1.0516**	**1.1725**
γ_0	Mean/Bias	0.25	0.2658/0.0158	0.2664/0.0164	0.2603/0.0103
	SD		0.3582	0.3221	0.2941
	MSE		0.1286	0.1040	0.0866
	Relative Efficiency		**1.0000**	**1.2369**	**1.4833**
γ_1	Mean/Bias	-0.1	-0.1165/-0.0165	-0.1110/-0.0110	-0.1079/-0.0079
	SD		0.4554	0.3962	0.3723
	MSE		0.2076	0.1571	0.1387
	Relative Efficiency		**1.0000**	**1.3212**	**1.4956**
γ_2	Mean/Bias	0.5	0.5169/0.0169	0.5077/0.0077	0.5073/0.0073
	SD		0.4356	0.4061	0.3779
	MSE		0.1900	0.1650	0.1429
	Relative Efficiency		**1.0000**	**1.1506**	**1.3287**
γ_3	Mean/Bias	-0.1	-0.0959/0.0041	-0.0966/0.0034	-0.0914/0.0086
	SD		0.3739	0.3301	0.3059
	MSE		0.1398	0.1090	0.0937
	Relative Efficiency		**1.0000**	**1.2832**	**1.4937**

Figure 4.22: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

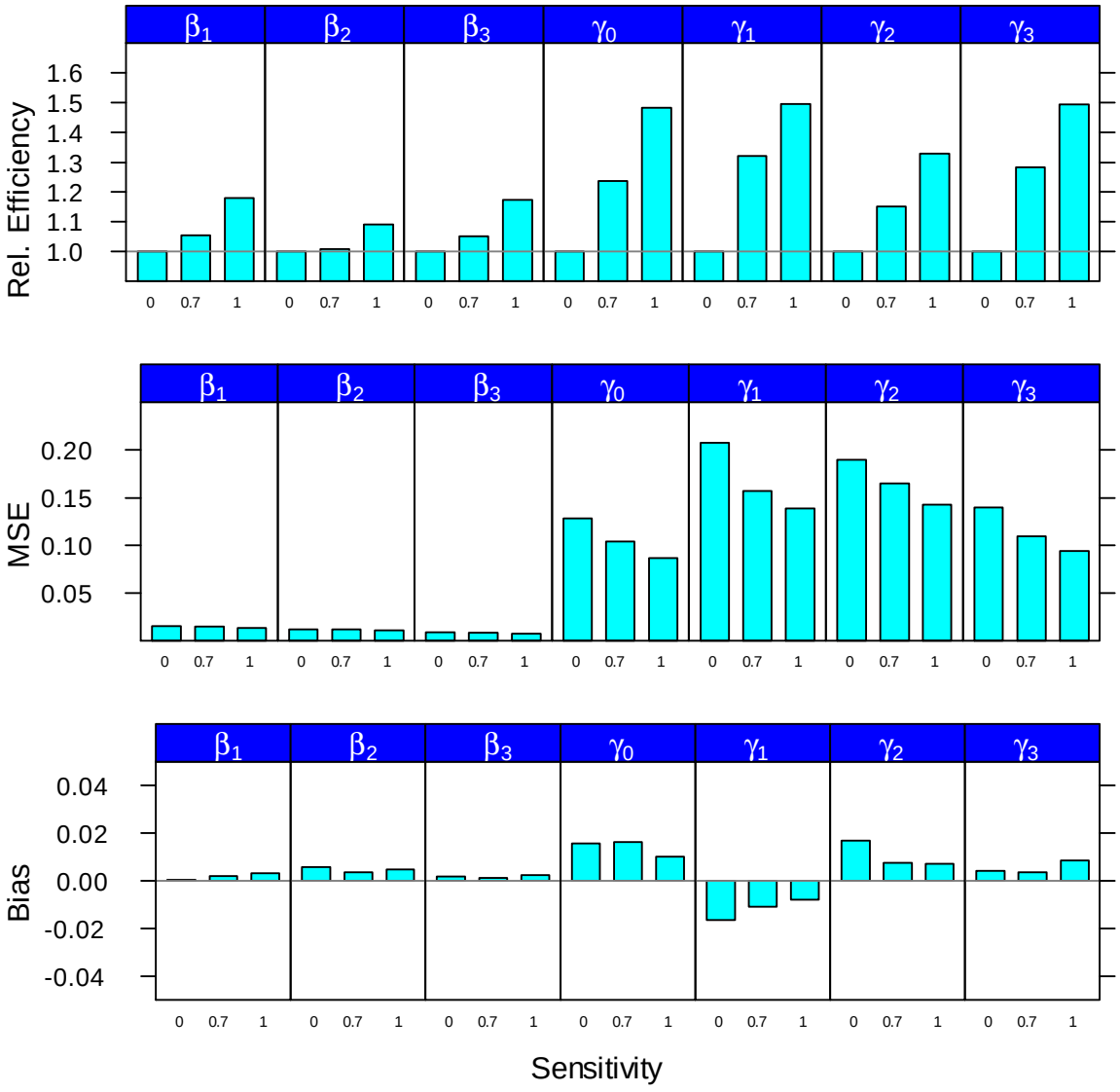


Table 4.25: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2004/0.0004	0.2032/0.0032	0.2025/0.0025
	SD		0.1245	0.1201	0.1157
	MSE		0.0155	0.0144	0.0134
	Relative Efficiency		**1.0000**	**1.0742**	**1.1579**
β_2	Mean/Bias	-0.3	-0.2942/0.0058	-0.2947/0.0053	-0.2955/0.0045
	SD		0.1085	0.1074	0.1049
	MSE		0.0118	0.0116	0.0110
	Relative Efficiency		**1.0000**	**1.0205**	**1.0681**
β_3	Mean/Bias	0.1	0.1018/0.0018	0.1015/0.0015	0.1016/0.0016
	SD		0.0925	0.0913	0.0866
	MSE		0.0086	0.0083	0.0075
	Relative Efficiency		**1.0000**	**1.0273**	**1.1413**
γ_0	Mean/Bias	0.25	0.2658/0.0158	0.2853/0.0353	0.2958/0.0458
	SD		0.3582	0.3344	0.3051
	MSE		0.1286	0.1131	0.0952
	Relative Efficiency		**1.0000**	**1.1477**	**1.3783**
γ_1	Mean/Bias	-0.1	-0.1165/-0.0165	-0.1024/-0.0024	-0.0987/0.0013
	SD		0.4554	0.4066	0.3806
	MSE		0.2076	0.1653	0.1449
	Relative Efficiency		**1.0000**	**1.2541**	**1.4312**
γ_2	Mean/Bias	0.5	0.5169/0.0169	0.5161/0.0161	0.5081/0.0081
	SD		0.4356	0.4146	0.3895
	MSE		0.1900	0.1722	0.1518
	Relative Efficiency		**1.0000**	**1.1037**	**1.2508**
γ_3	Mean/Bias	-0.1	-0.0959/0.0041	-0.1000/0.0000	-0.0927/0.0073
	SD		0.3739	0.3457	0.3172
	MSE		0.1398	0.1195	0.1007
	Relative Efficiency		**1.0000**	**1.1697**	**1.3894**

Figure 4.23: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

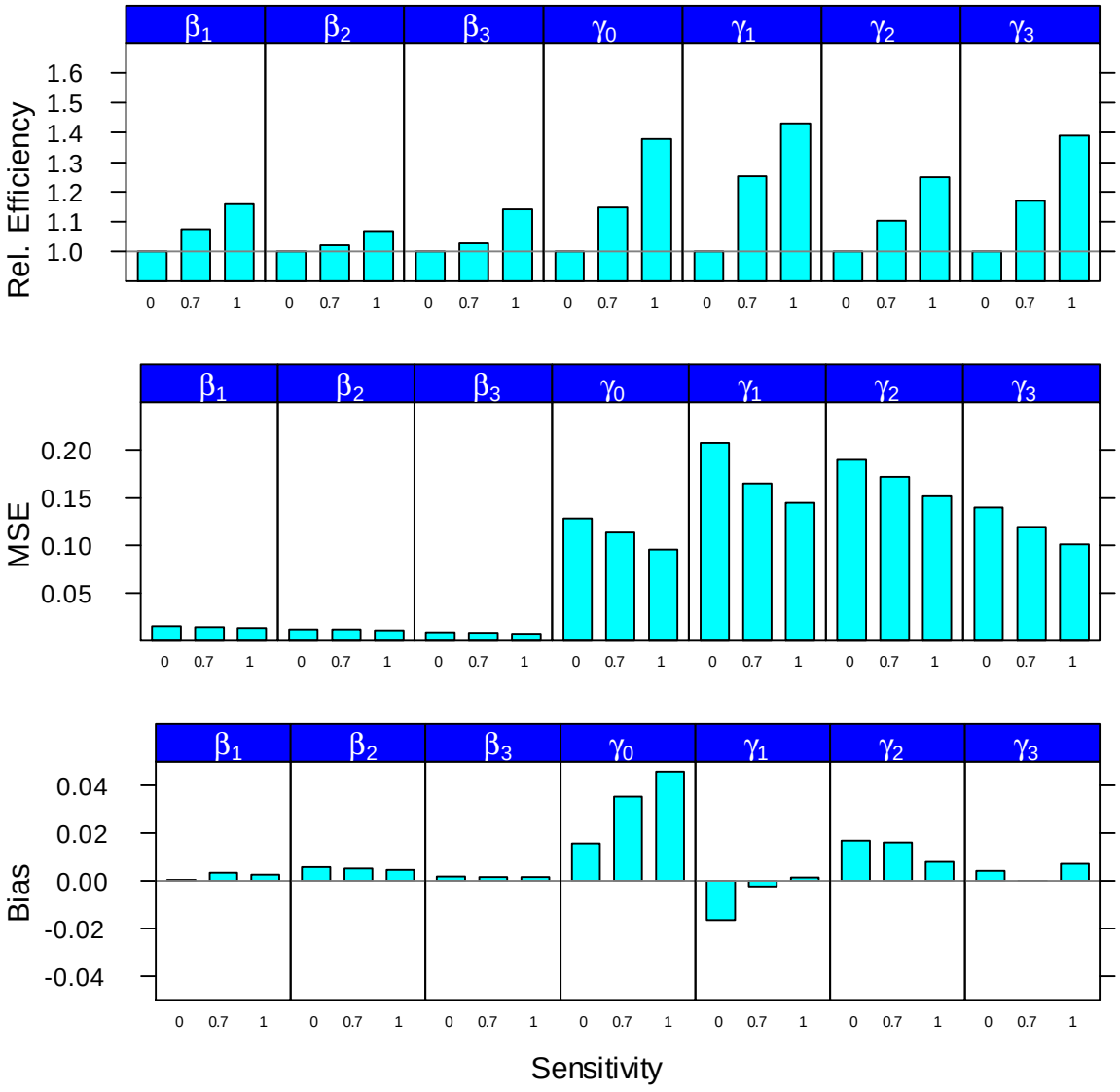


Table 4.26: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

	Statistics	True Parameter	No Known Cure Info	p_0	
				0.7	1
β_1	Mean/Bias	0.2	0.2001/0.0001	0.2058/0.0058	0.2029/0.0029
	SD		0.1398	0.1215	0.1057
	MSE		0.0195	0.0148	0.0112
	Relative Efficiency		**1.0000**	**1.3240**	**1.7482**
β_2	Mean/Bias	-0.3	-0.2965/0.0035	-0.3000/0.0000	-0.3010/-0.0010
	SD		0.1073	0.1015	0.0930
	MSE		0.0115	0.0103	0.0086
	Relative Efficiency		**1.0000**	**1.1167**	**1.3312**
β_3	Mean/Bias	0.1	0.1037/0.0037	0.1024/0.0024	0.1024/0.0024
	SD		0.0974	0.0861	0.0797
	MSE		0.0095	0.0074	0.0064
	Relative Efficiency		**1.0000**	**1.2789**	**1.4921**
γ_0	Mean/Bias	0.25	0.2802/0.0302	0.2646/0.0146	0.2611/0.0111
	SD		0.4856	0.3397	0.2822
	MSE		0.2368	0.1156	0.0797
	Relative Efficiency		**1.0000**	**2.0437**	**2.9622**
γ_1	Mean/Bias	-0.1	-0.0233/0.0767	-0.0910/0.0090	-0.1002/-0.0002
	SD		1.2217	0.4324	0.3526
	MSE		1.4984	0.1871	0.1243
	Relative Efficiency		**1.0000**	**7.9820**	**12.0050**
γ_2	Mean/Bias	0.5	0.5030/0.0030	0.5019/0.0019	0.4931/-0.0069
	SD		0.5875	0.4248	0.3701
	MSE		0.3452	0.1805	0.1370
	Relative Efficiency		**1.0000**	**1.9128**	**2.5206**
γ_3	Mean/Bias	-0.1	-0.1071/-0.0071	-0.1239/-0.0239	-0.1226/-0.0226
	SD		0.4966	0.3463	0.2964
	MSE		0.2466	0.1205	0.0884
	Relative Efficiency		**1.0000**	**2.0562**	**2.8071**

Figure 4.24: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

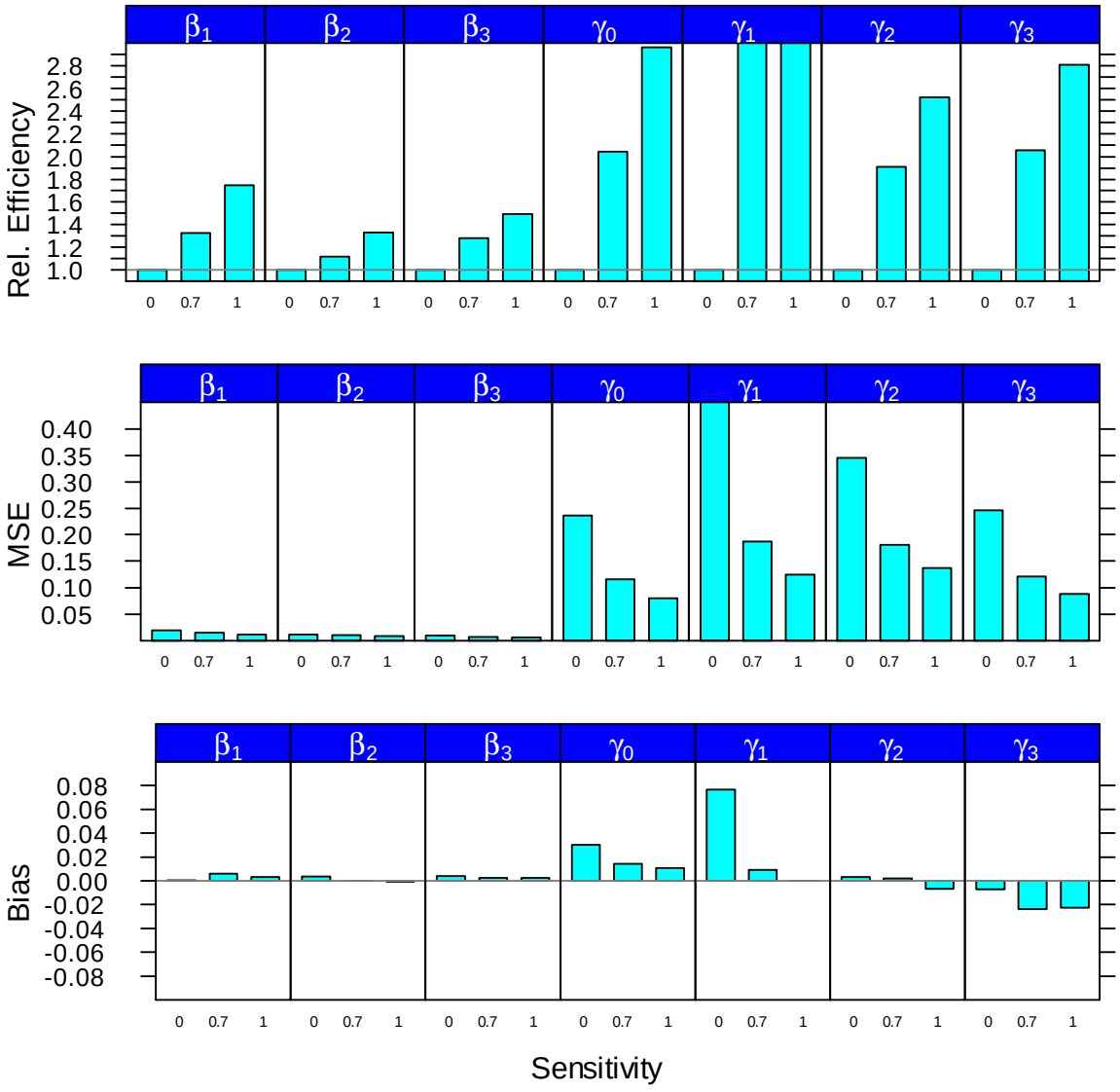
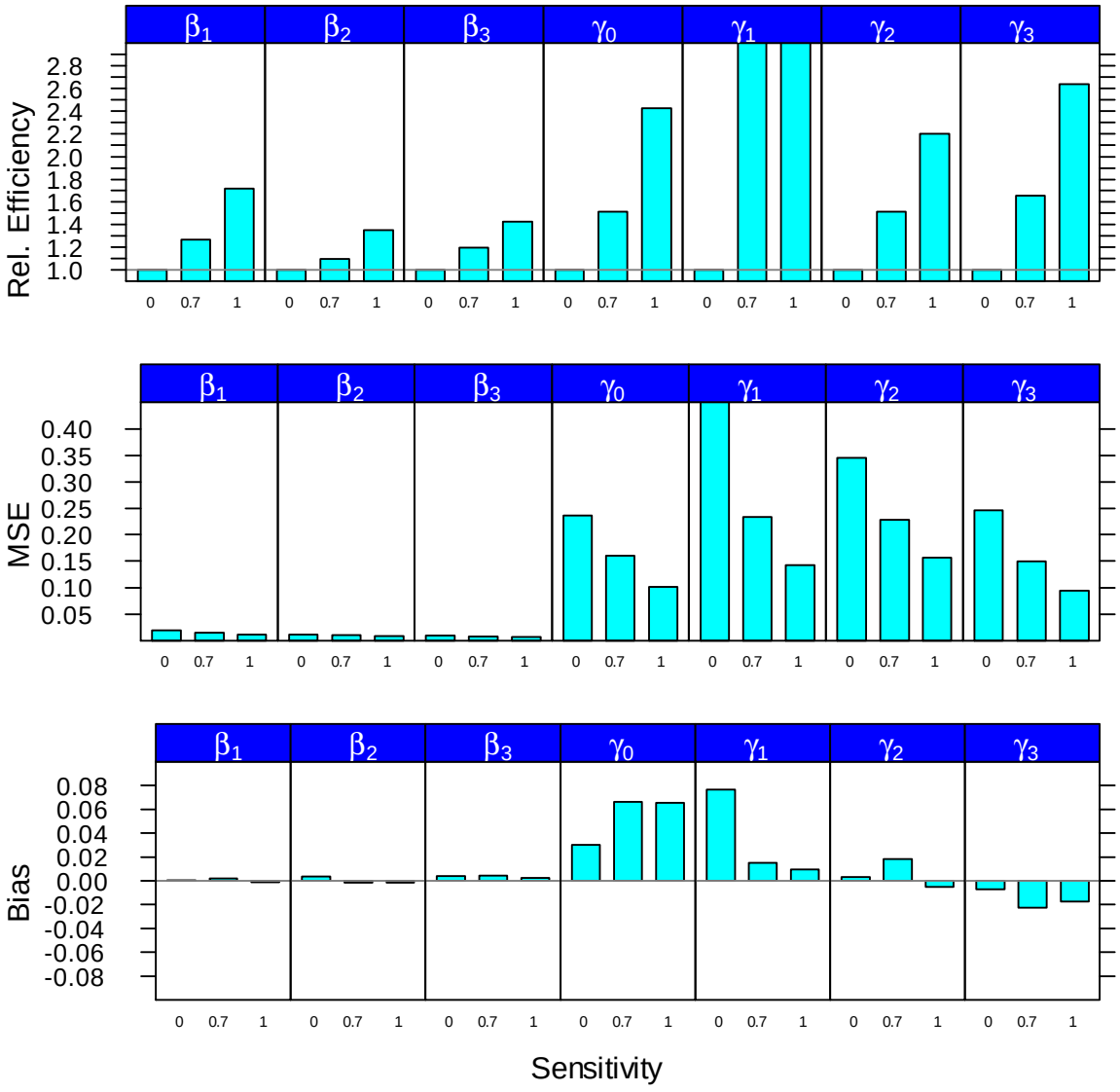


Table 4.27: AFT Simulation Results for 100% Specificity, and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)

		True	No Known	p_0	
	Statistics	Parameter	Cure Info	0.7	1
β_1	Mean/Bias	0.2	0.2001/0.0001	0.2018/0.0018	0.1986/-0.0014
	SD		0.1398	0.1241	0.1066
	MSE		0.0195	0.0154	0.0114
	Relative Efficiency		**1.0000**	**1.2672**	**1.7173**
β_2	Mean/Bias	-0.3	-0.2965/0.0035	-0.3016/-0.0016	-0.3019/-0.0019
	SD		0.1073	0.1024	0.0924
	MSE		0.0115	0.0105	0.0085
	Relative Efficiency		**1.0000**	**1.0983**	**1.3481**
β_3	Mean/Bias	0.1	0.1037/0.0037	0.1042/0.0042	0.1022/0.0022
	SD		0.0974	0.0890	0.0815
	MSE		0.0095	0.0079	0.0067
	Relative Efficiency		**1.0000**	**1.1959**	**1.4260**
γ_0	Mean/Bias	0.25	0.2802/0.0302	0.3163/0.0663	0.3154/0.0654
	SD		0.4856	0.3948	0.3119
	MSE		0.2368	0.1602	0.1015
	Relative Efficiency		**1.0000**	**1.5133**	**2.4249**
γ_1	Mean/Bias	-0.1	-0.0233/0.0767	-0.0847/0.0153	-0.0904/0.0096
	SD		1.2217	0.4838	0.3773
	MSE		1.4984	0.2343	0.1425
	Relative Efficiency		**1.0000**	**6.3762**	**10.4829**
γ_2	Mean/Bias	0.5	0.5030/0.0030	0.5186/0.0186	0.4949/-0.0051
	SD		0.5875	0.4778	0.3960
	MSE		0.3452	0.2286	0.1568
	Relative Efficiency		**1.0000**	**1.5121**	**2.2013**
γ_3	Mean/Bias	-0.1	-0.1071/-0.0071	-0.1226/-0.0226	-0.1176/-0.0176
	SD		0.4966	0.3861	0.3058
	MSE		0.2466	0.1496	0.0938
	Relative Efficiency		**1.0000**	**1.6545**	**2.6362**

Figure 4.25: Bias and Relative Efficiency for the AFT Simulation Results of 100% Specificity ($p_1 = 0$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity (Using Zhang and Peng's Estimation Method)



4.4.3 Comparison of estimation of proposed models using Zhang and Peng's method and Li and Taylor's method

Comparing the simulation results between using Zhang and Peng's method (2007) and using Li and Taylor's method (2002), first we consider the differences in bias when baseline Weibull hazard function with ($h = \frac{1}{3}$, $k = 4$) is used. Point estimators are consistent for all the cases considered using Zhang and Peng's method, while Li and Taylor's method produced non-consistent estimates when ($h = \frac{1}{3}$, $k = 4$). Zhang and Peng (2007) commented on the consistency issue of Li and Taylor's method. In our simulation case, this consistency issue arose when the censoring rate is high.

Comparing efficiency gains of the other three baseline function cases, Zhang and Peng's method showed higher efficiency gain when ($h = 2$, $k = 2$). The efficiency gains are similar for the other two cases.

4.5 Comparison of the Results between Semi-Parametric PH Cure Rate Models and Semi-Parametric AFT Cure Rate Models

For our evaluation presented in the previous sections, we observed efficiency gains when applying the proposed models with either semi-parametric AFT or PH cure rate models. When using Zhang and Peng's method for the parameter estimation, the magnitude and pattern of the efficiency gain from the AFT simulations is in line with the results from the PH simulations.

Chapter 5

Application to Pediatric Bone Data

This is a clinical trial that retrospectively reviewed 157 children's charts to identify the incidence of premature physeal closure (PPC) following physeal fractures of the distal end of the tibia (Leary et al, 2009). Sixteen out of the 157 children were identified as having PPC. Other children were considered to be cured if symmetric Harris growth arrest line was observed or closure of the growth plate was seen radiographically. Ninety-six children were considered as cured. A Kaplan-Meier curve of the time to PPC together with 95% confidence interval is shown in Figure 5.1.

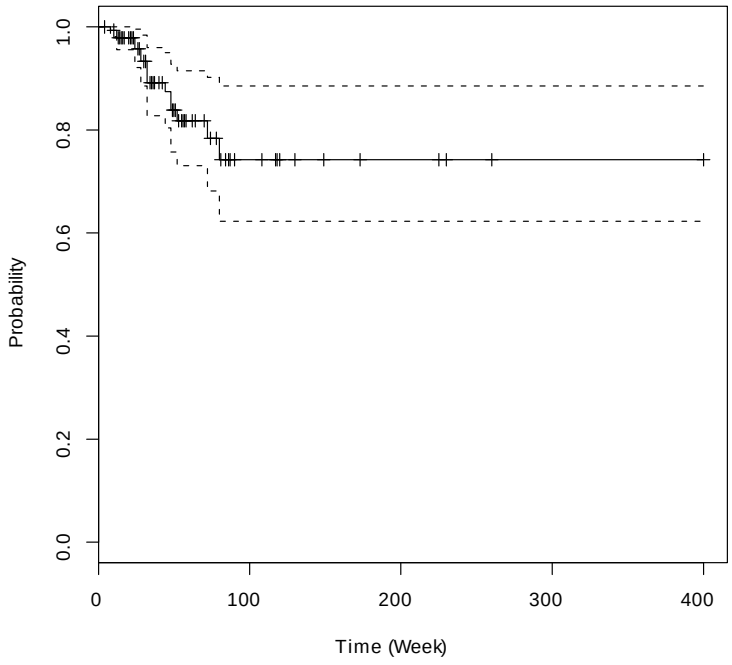


Figure 5.1: Time to Premature Physeal Closure

Since there is clear indication of cure in this data, cure rate model is an appropriate model to the analysis. The cure of a subject could be indicated by symmetric Harris growth arrest line or closure of the growth plate. Since the cure indicator in this case is definitive, it is treated as a diagnostic procedure with known 100% sensitivity and specificity. For subjects who did not have PPC event noted, and did not have cure indictor result available, they are treated as having missing diagnostic procedure result. The factor of treatment methods (Cast, or non-Cast) and gender effects were included in the survival portion and the cured portion of the model.

Table 5.1 compares the results of applying semi-parametric PH cure rate model with and without diagnostic information. Both models provided similar sign and magnitude of point estimators, while the standard error from the model with diagnostic information incorporated are much smaller. This observation is consistent with our simulation results. More efficient estimates lead to more powerful test and might lead to different conclusion. As shown in the table, the 2-sided P-value for logistic intercept is 0.6850 when diagnostic information is not incorporated in the model; this is changed to a significant P-value 0.0214 when diagnostic information is incorporated. This is consistent to the pattern of the data: we have 96 subjects that were cured, so the incidence part of the model can not be ignored. Similar observation is also noted for Cast factor in logistic portion: the 2-sided P-value is 0.9142 without diagnostic information included, and marginally significant (0.0841) with diagnostic information included.

Table 5.1: Application of PH Cure Rate Model to Pediatric Bone Data

	Effect	Model without Diagnostic Information			Model with Diagnostic Information		
		Mean	SE	p	Mean	SE	p
Survival Portion	Male	-0.479	3.245	0.8827	-0.638	2.134	0.7649
	Cast	-0.447	2.602	0.8636	-0.083	2.066	0.9678
Logistic Portion	Intercept	-0.536	1.321	**0.6850**	-1.189	0.517	**0.0214**
	Male	-0.610	1.506	0.6856	-0.337	0.475	0.4786
	Cast	-0.155	1.438	**0.9142**	-0.887	0.514	**0.0841**

Table 5.2 compares the result of applying semi-parametric AFT cure rate model

with and without diagnostic information included. The parameter estimation in the AFT model was performed using Zhang and Peng's approach (2007). During the bootstrapping step, if the point estimate has absolute value over 1000, the bootstrap sample is treated as not converged. The survival portion shows different signs, while the P-values in the survival part show similar non-significant conclusions. The results for the incidence part from both models show similar sign and magnitude of point estimators, while the standard error from the model with diagnostic information incorporated are much smaller. Similar to the PH cure rate model, the 2-sided P-value for logistic intercept is 0.7230 when diagnostic information is not incorporated in the model; this is changed to a significant P-value 0.0197 when diagnostic information is incorporated. P-value for Cast factor in logistic portion is also much smaller when diagnostic information is included. The incidence part of point estimates is similar from the AFT model compared to that from the PH model. The latency part of point estimates is different, although the same non-significant conclusion can be made.

Table 5.2: Application of AFT Cure Rate Model to Pediatric Bone Data

	Effect	Model without Diagnostic Information			Model with Diagnostic Information		
		Mean	SE	p	Mean	SE	p
Survival Portion	Male	-0.000	0.784	> 0.9999	0.134	0.509	0.7930
	Cast	-0.223	0.731	0.7603	0.151	0.453	0.7393
Logistic Portion	Intercept	-0.650	1.834	**0.7230**	-1.189	0.508	**0.0197**
	Male	-0.718	3.791	0.8498	-0.344	0.475	0.4693
	Cast	-0.074	4.382	**0.9865**	-0.896	0.510	**0.0787**

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Traditional cure rate models assume the status of uncured and cured in the censored set can not be distinguished. With the inclusion of additional diagnostic information, the status of cured and uncured can be partially distinguished to the extent of certain sensitivity and specificity. A model that includes this additional information has the potential of improving the efficiency of the estimation. In this thesis, a novel extension of existing cure rate models is proposed. This extension can be applied to parametric, semi-parametric, or non-parametric cure rate models. Both AFT and PH models can be handled by this model extension.

Theoretical justifications are provided for some cases with no covariates, or with simple covariates. Additionally, simulations are performed for both semi-parametric AFT and PH cure rate models, for more complicate scenarios. The theoretical justifications show that the efficiency gain increases with higher sensitivity and specificity of the diagnostic procedure under special cases. The simulations confirm this result under more complicate scenarios. For PH simulations, all the biases are small, MSEs are smaller when additional diagnostic procedures are incorporated, and relative efficiencies are greater than one when additional diagnostic procedures are incorporated. Higher sensitivity and specificity are associated with smaller MSE and larger relative efficiency. AFT simulations using Zhang and Peng's method (2007) for parameter estimation show similar results and extent of efficiency gains. AFT simulations using Li and Taylor's method (2002) show efficiency gain under some test cases, but fail to provide consistent estimators under some other. For both PH simulations and AFT simulations using Zhang and Peng's method for parameter estimation, larger efficiency

gain is noted when the rate of censoring is high. When the censoring rate is high, the set of subjects with uncured and cured status undetermined is larger. So, bringing additional diagnostic information provides larger extent of efficiency gain.

6.2 Future Work

6.2.1 Impact of Tail Completion Methods to Parameter Estimates and Relative Efficiency

Zero-tail completion is a very convenient and widely used tail completion methods in the cure rate model literature. On the other hand, by using zero-completion, we are implicitly assuming cured status for censored observations beyond the last event time. This can potentially biased the result, and may conflict with the indicator of cured status based on a diagnostic procedure. In this thesis, we explored the exponential tail completion in the case of AFT simulation, without much improvement noted. To further quantify the impact of tail completion methods on parameter estimates and relative efficiency, further simulation studies will be conducted, and theoretical impact of different tail completion scheme will be explored.

6.2.2 Variance Estimation Based on Louis Method or Other Non-Iterative Method

By the use of EM algorithm, we can separate the complete likelihood into multiple parts, and maximize each parts separately. This approach avoids many of the issues that are related to joint estimation of all parameters. Because information matrix is not directly available from the EM iteration steps, many of current applications use bootstrap methods to obtain variance estimation. Louis Method (1982) provides an alternative to the bootstrap methods. The property of the variance estimates from these two methods will be studied and compared. Other non-iterative methods will also be considered.

6.2.3 Consistency and Asymptotic Distribution of the Estimator, General Case

Further theoretical justification for the consistency of the estimator needs to be established, and the asymptotic distribution of the estimator needs to be explored. Fang, Li, and Sun (2005) showed the consistency results for cure rate models with semi-parametric PH specification. Extension of their results to the proposed model extension with semi-parametric PH or AFT specification will be explored.

6.2.4 Asymptotic Properties of the Estimator

Improvement on relative efficiency is an important feature of the proposed model. If the proposed model extension works as expected, to design a new experiment or trial using cure rate model, we have the possibility to increase efficiency of the estimator by introducing a diagnostic procedure. A smaller sample size could be achieved with this improved efficiency. Evaluation for relative efficiency under strong parametric assumption was performed in Section 4.1. This relative efficiency gains under less stringent parametric or semi-parametric assumptions are also potential areas for further study.

Appendix A

Additional Results for the Evaluation of Semi-Parametric PH Cure Rate Models

A.1 Additional Results using Weibull distribution with parameter $h = 1$ and $k = 2$

Table A.1: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1946/-0.0054	0.1942/-0.0058	0.1935/-0.0065	0.1943/-0.0057	0.1948/-0.0052
	SD		0.2872	0.2872	0.2882	0.2846	0.2764
	MSE		0.0825	0.0825	0.0831	0.0810	0.0764
	Relative Efficiency		**1.0000**	**1.0002**	**0.9930**	**1.0186**	**1.0802**
β_2	Mean/Bias	-0.3	-0.3115/-0.0115	-0.3117/-0.0117	-0.3115/-0.0115	-0.3120/-0.0120	-0.3123/-0.0123
	SD		0.2728	0.2722	0.2718	0.2715	0.2623
	MSE		0.0745	0.0743	0.0740	0.0738	0.0690
	Relative Efficiency		**1.0000**	**1.0038**	**1.0069**	**1.0095**	**1.0810**
β_3	Mean/Bias	0.1	0.0989/-0.0011	0.0990/-0.0010	0.0983/-0.0017	0.0980/-0.0020	0.0989/-0.0011
	SD		0.2300	0.2298	0.2282	0.2277	0.2176
	MSE		0.0529	0.0528	0.0521	0.0518	0.0474
	Relative Efficiency		**1.0000**	**1.0015**	**1.0159**	**1.0205**	**1.1172**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2577/0.0077	0.2573/0.0073	0.2535/0.0035	0.2465/-0.0035
	SD		0.3251	0.3240	0.3197	0.3137	0.2984
	MSE		0.1058	0.1050	0.1022	0.0984	0.0891
	Relative Efficiency		**1.0000**	**1.0071**	**1.0343**	**1.0741**	**1.1866**
γ_1	Mean/Bias	-0.1	-0.0944/0.0056	-0.0941/0.0059	-0.0935/0.0065	-0.0922/0.0078	-0.0898/0.0102
	SD		0.4009	0.4004	0.3955	0.3906	0.3685
	MSE		0.1607	0.1603	0.1564	0.1526	0.1359
	Relative Efficiency		**1.0000**	**1.0027**	**1.0275**	**1.0534**	**1.1835**
γ_2	Mean/Bias	0.5	0.5197/0.0197	0.5204/0.0204	0.5208/0.0208	0.5227/0.0227	0.5230/0.0230
	SD		0.4305	0.4300	0.4224	0.4103	0.3839
	MSE		0.1857	0.1853	0.1788	0.1689	0.1479
	Relative Efficiency		**1.0000**	**1.0025**	**1.0389**	**1.1008**	**1.2576**
γ_3	Mean/Bias	-0.1	-0.1112/-0.0112	-0.1108/-0.0108	-0.1095/-0.0095	-0.1072/-0.0072	-0.1036/-0.0036
	SD		0.3304	0.3302	0.3287	0.3231	0.3085
	MSE		0.1093	0.1091	0.1082	0.1045	0.0952
	Relative Efficiency		**1.0000**	**1.0011**	**1.0100**	**1.0452**	**1.1471**

Figure A.1: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity

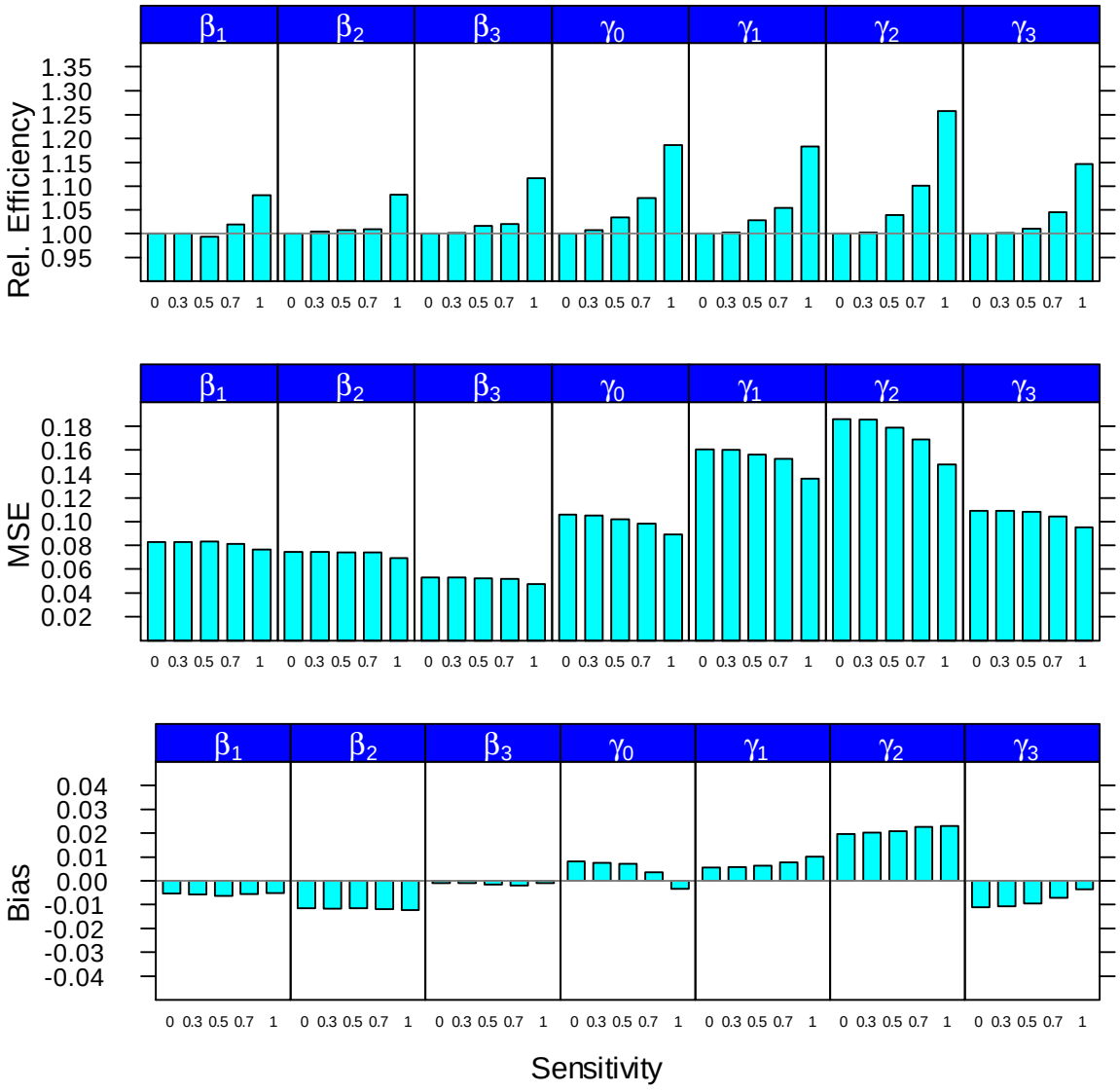


Table A.2: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1946/-0.0054	0.1938/-0.0062	0.1945/-0.0055	0.1954/-0.0046	0.1944/-0.0056
	SD		0.2872	0.2883	0.2897	0.2870	0.2760
	MSE		0.0825	0.0832	0.0840	0.0824	0.0762
	Relative Efficiency		**1.0000**	**0.9924**	**0.9830**	**1.0012**	**1.0831**
β_2	Mean/Bias	-0.3	-0.3115/-0.0115	-0.3123/-0.0123	-0.3106/-0.0106	-0.3130/-0.0130	-0.3136/-0.0136
	SD		0.2728	0.2743	0.2731	0.2731	0.2635
	MSE		0.0745	0.0754	0.0747	0.0747	0.0696
	Relative Efficiency		**1.0000**	**0.9890**	**0.9973**	**0.9978**	**1.0716**
β_3	Mean/Bias	0.1	0.0989/-0.0011	0.0998/-0.0002	0.0989/-0.0011	0.0988/-0.0012	0.0990/-0.0010
	SD		0.2300	0.2308	0.2280	0.2287	0.2192
	MSE		0.0529	0.0533	0.0520	0.0523	0.0480
	Relative Efficiency		**1.0000**	**0.9930**	**1.0174**	**1.0114**	**1.1013**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2604/0.0104	0.2617/0.0117	0.2577/0.0077	0.2563/0.0063
	SD		0.3251	0.3267	0.3230	0.3193	0.3038
	MSE		0.1058	0.1068	0.1044	0.1020	0.0923
	Relative Efficiency		**1.0000**	**0.9904**	**1.0133**	**1.0368**	**1.1453**
γ_1	Mean/Bias	-0.1	-0.0944/0.0056	-0.0955/0.0045	-0.0968/0.0032	-0.0948/0.0052	-0.0903/0.0097
	SD		0.4009	0.4017	0.3973	0.3967	0.3720
	MSE		0.1607	0.1614	0.1579	0.1574	0.1385
	Relative Efficiency		**1.0000**	**0.9958**	**1.0182**	**1.0214**	**1.1615**
γ_2	Mean/Bias	0.5	0.5197/0.0197	0.5241/0.0241	0.5206/0.0206	0.5314/0.0314	0.5301/0.0301
	SD		0.4305	0.4343	0.4278	0.4145	0.3915
	MSE		0.1857	0.1892	0.1835	0.1728	0.1542
	Relative Efficiency		**1.0000**	**0.9826**	**1.0125**	**1.0785**	**1.2090**
γ_3	Mean/Bias	-0.1	-0.1112/-0.0112	-0.1145/-0.0145	-0.1118/-0.0118	-0.1075/-0.0075	-0.1076/-0.0076
	SD		0.3304	0.3327	0.3342	0.3268	0.3117
	MSE		0.1093	0.1109	0.1118	0.1068	0.0972
	Relative Efficiency		**1.0000**	**0.9859**	**0.9772**	**1.0221**	**1.1235**

Figure A.2: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity

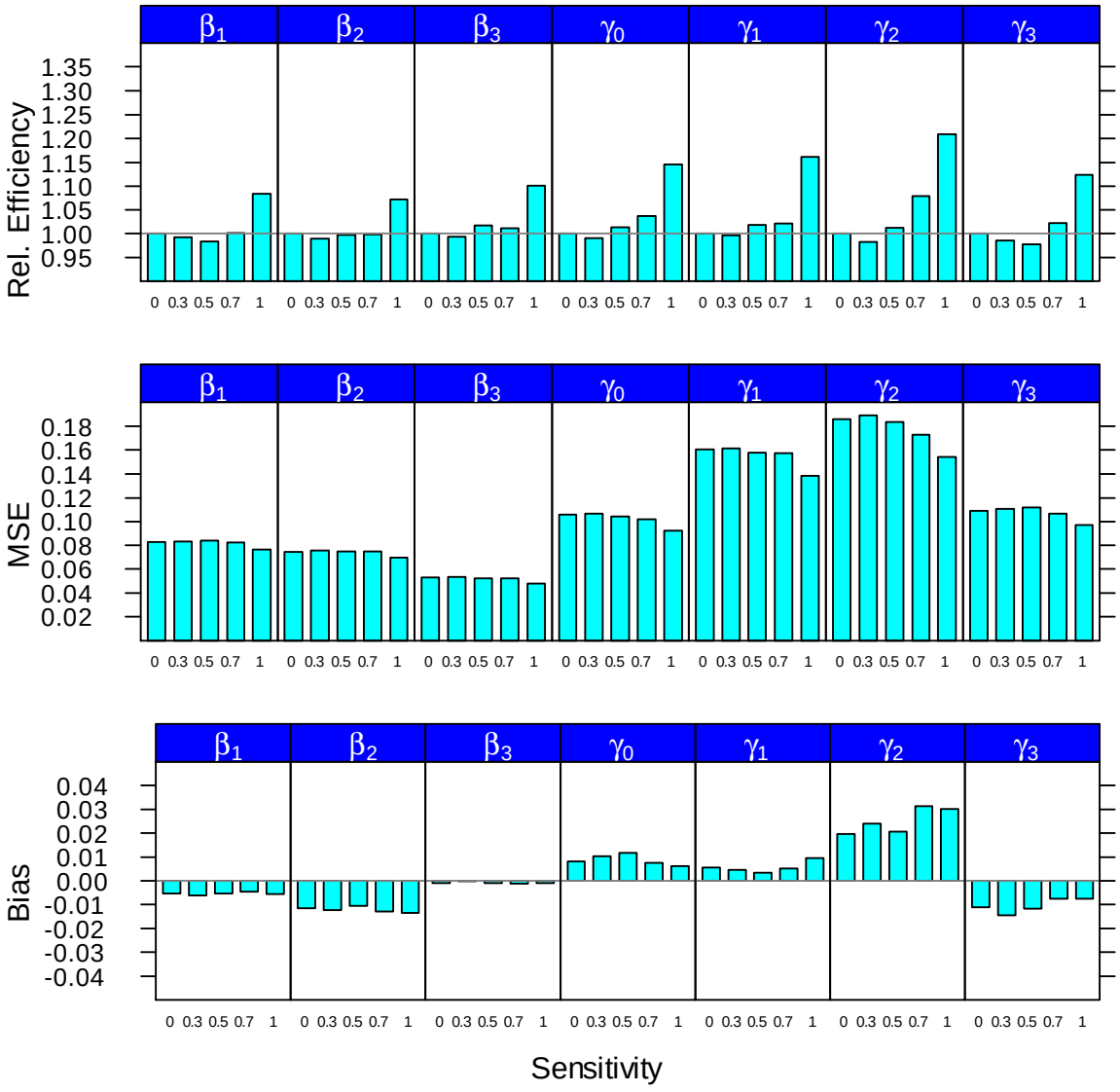


Table A.3: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1946/-0.0054	0.1937/-0.0063	0.1932/-0.0068	0.1942/-0.0058	0.1940/-0.0060
	SD		0.2872	0.2871	0.2886	0.2840	0.2761
	MSE		0.0825	0.0825	0.0833	0.0807	0.0763
	Relative Efficiency		**1.0000**	**1.0007**	**0.9904**	**1.0225**	**1.0822**
β_2	Mean/Bias	-0.3	-0.3115/-0.0115	-0.3117/-0.0117	-0.3112/-0.0112	-0.3116/-0.0116	-0.3119/-0.0119
	SD		0.2728	0.2715	0.2713	0.2708	0.2611
	MSE		0.0745	0.0739	0.0737	0.0735	0.0683
	Relative Efficiency		**1.0000**	**1.0091**	**1.0108**	**1.0145**	**1.0909**
β_3	Mean/Bias	0.1	0.0989/-0.0011	0.0991/-0.0009	0.0981/-0.0019	0.0978/-0.0022	0.0995/-0.0005
	SD		0.2300	0.2295	0.2273	0.2268	0.2167
	MSE		0.0529	0.0527	0.0517	0.0515	0.0469
	Relative Efficiency		**1.0000**	**1.0046**	**1.0240**	**1.0282**	**1.1269**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2574/0.0074	0.2575/0.0075	0.2530/0.0030	0.2467/-0.0033
	SD		0.3251	0.3223	0.3172	0.3111	0.2963
	MSE		0.1058	0.1039	0.1007	0.0968	0.0878
	Relative Efficiency		**1.0000**	**1.0176**	**1.0506**	**1.0920**	**1.2041**
γ_1	Mean/Bias	-0.1	-0.0944/0.0056	-0.0944/0.0056	-0.0942/0.0058	-0.0927/0.0073	-0.0908/0.0092
	SD		0.4009	0.3988	0.3919	0.3866	0.3647
	MSE		0.1607	0.1591	0.1536	0.1495	0.1331
	Relative Efficiency		**1.0000**	**1.0106**	**1.0462**	**1.0755**	**1.2081**
γ_2	Mean/Bias	0.5	0.5197/0.0197	0.5190/0.0190	0.5176/0.0176	0.5188/0.0188	0.5172/0.0172
	SD		0.4305	0.4274	0.4161	0.4018	0.3762
	MSE		0.1857	0.1831	0.1734	0.1618	0.1418
	Relative Efficiency		**1.0000**	**1.0144**	**1.0706**	**1.1481**	**1.3094**
γ_3	Mean/Bias	-0.1	-0.1112/-0.0112	-0.1100/-0.0100	-0.1083/-0.0083	-0.1056/-0.0056	-0.1025/-0.0025
	SD		0.3304	0.3275	0.3241	0.3162	0.2994
	MSE		0.1093	0.1074	0.1051	0.1000	0.0896
	Relative Efficiency		**1.0000**	**1.0173**	**1.0388**	**1.0913**	**1.2179**

Figure A.3: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Known Sensitivity and Specificity

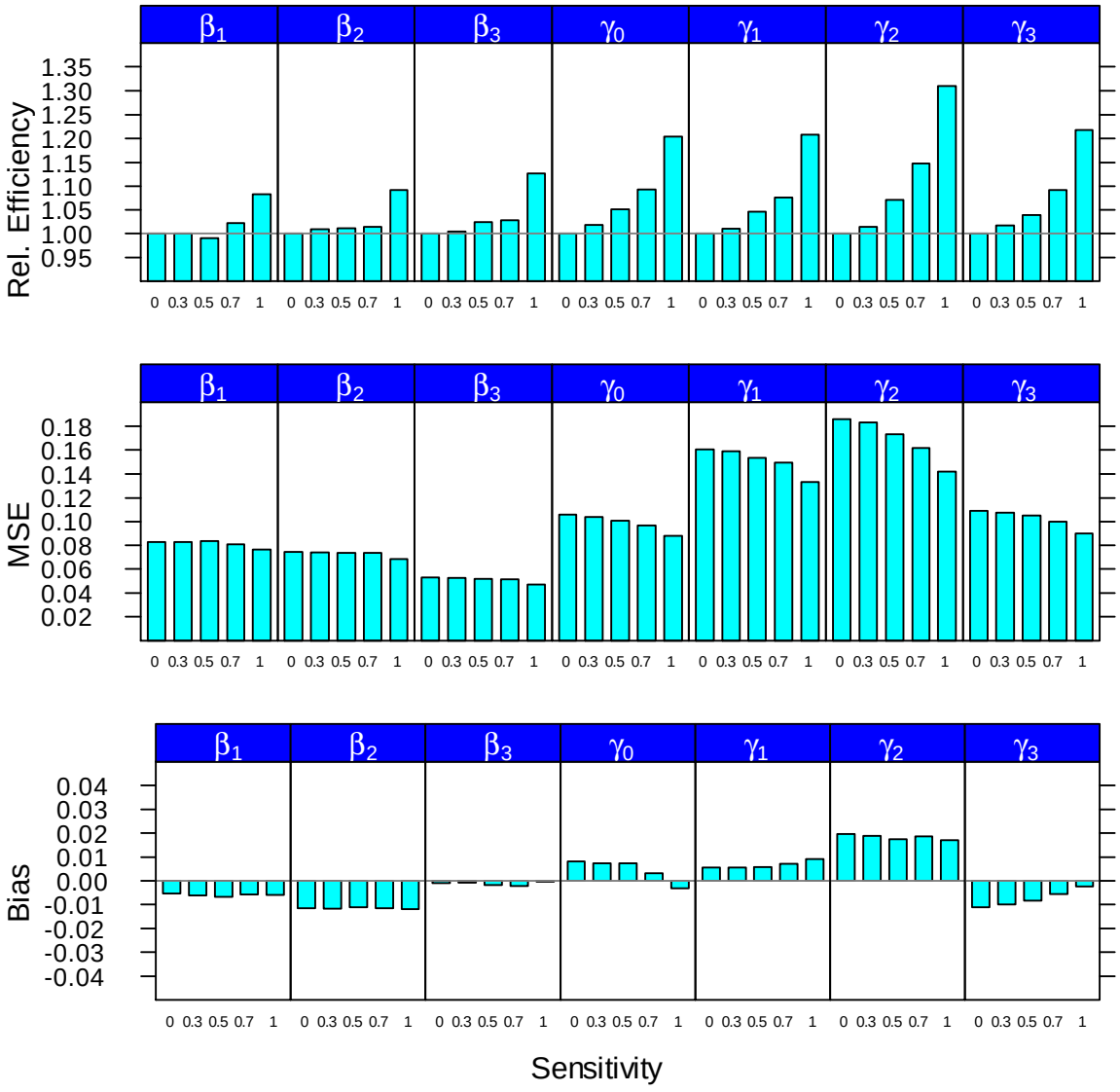
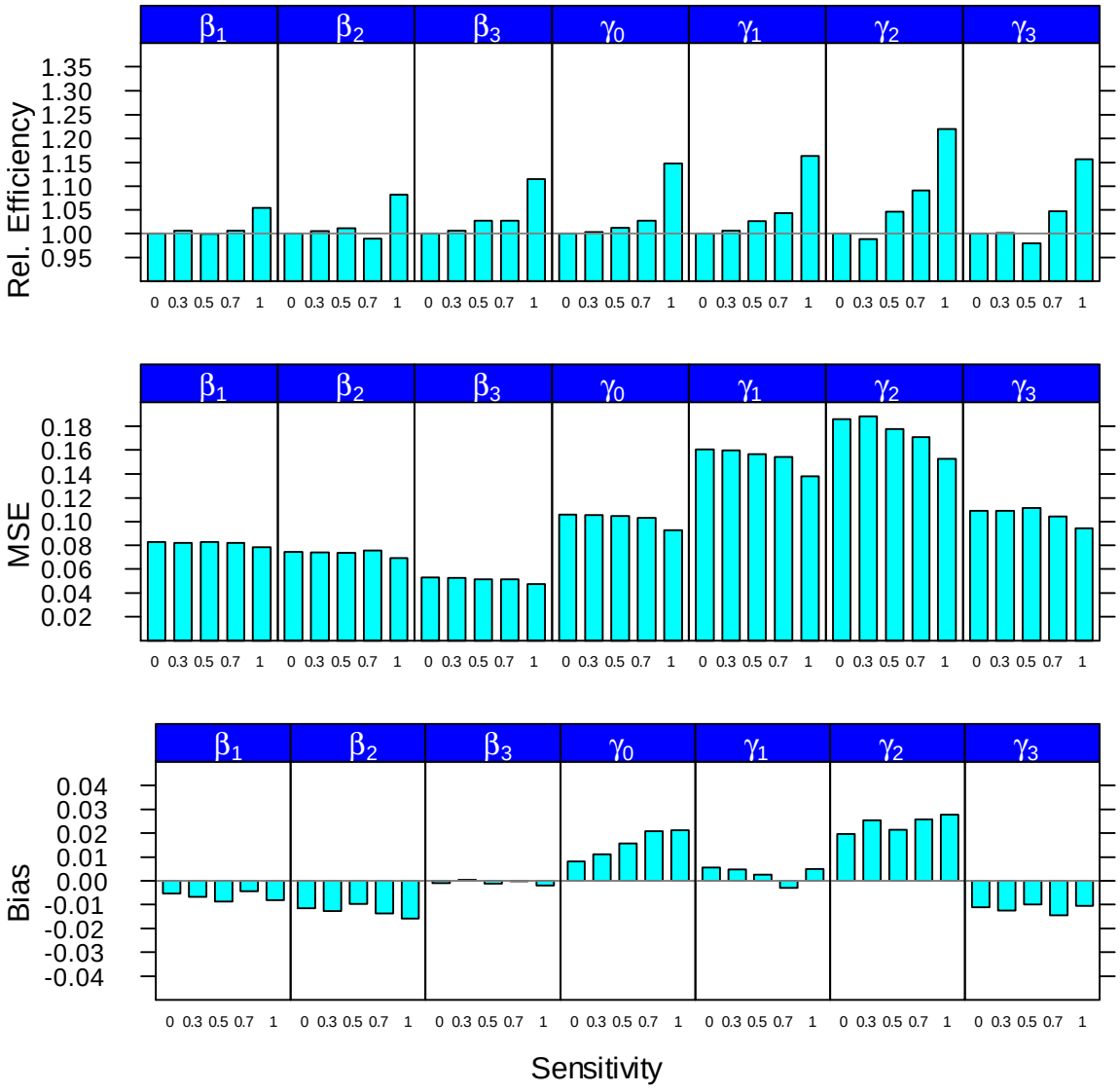


Table A.4: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.1946/-0.0054	0.1932/-0.0068	0.1913/-0.0087	0.1956/-0.0044	0.1918/-0.0082
	SD		0.2872	0.2863	0.2873	0.2864	0.2798
	MSE		0.0825	0.0820	0.0826	0.0820	0.0783
	Relative Efficiency		**1.0000**	**1.0064**	**0.9993**	**1.0058**	**1.0539**
β_2	Mean/Bias	-0.3	-0.3115/-0.0115	-0.3128/-0.0128	-0.3098/-0.0098	-0.3138/-0.0138	-0.3158/-0.0158
	SD		0.2728	0.2720	0.2713	0.2742	0.2622
	MSE		0.0745	0.0741	0.0737	0.0754	0.0690
	Relative Efficiency		**1.0000**	**1.0056**	**1.0111**	**0.9898**	**1.0818**
β_3	Mean/Bias	0.1	0.0989/-0.0011	0.1003/0.0003	0.0988/-0.0012	0.0997/-0.0003	0.0979/-0.0021
	SD		0.2300	0.2293	0.2270	0.2270	0.2178
	MSE		0.0529	0.0526	0.0515	0.0515	0.0475
	Relative Efficiency		**1.0000**	**1.0061**	**1.0271**	**1.0266**	**1.1149**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2612/0.0112	0.2657/0.0157	0.2708/0.0208	0.2712/0.0212
	SD		0.3251	0.3246	0.3231	0.3208	0.3035
	MSE		0.1058	0.1055	0.1047	0.1033	0.0925
	Relative Efficiency		**1.0000**	**1.0030**	**1.0121**	**1.0271**	**1.1476**
γ_1	Mean/Bias	-0.1	-0.0944/0.0056	-0.0951/0.0049	-0.0975/0.0025	-0.1030/-0.0030	-0.0950/0.0050
	SD		0.4009	0.3998	0.3957	0.3926	0.3717
	MSE		0.1607	0.1598	0.1566	0.1542	0.1382
	Relative Efficiency		**1.0000**	**1.0057**	**1.0263**	**1.0425**	**1.1634**
γ_2	Mean/Bias	0.5	0.5197/0.0197	0.5254/0.0254	0.5215/0.0215	0.5258/0.0258	0.5278/0.0278
	SD		0.4305	0.4330	0.4210	0.4123	0.3898
	MSE		0.1857	0.1881	0.1777	0.1707	0.1527
	Relative Efficiency		**1.0000**	**0.9886**	**1.0458**	**1.0901**	**1.2197**
γ_3	Mean/Bias	-0.1	-0.1112/-0.0112	-0.1126/-0.0126	-0.1100/-0.0100	-0.1145/-0.0145	-0.1105/-0.0105
	SD		0.3304	0.3301	0.3338	0.3229	0.3071
	MSE		0.1093	0.1091	0.1115	0.1045	0.0945
	Relative Efficiency		**1.0000**	**1.0015**	**0.9798**	**1.0466**	**1.1569**

Figure A.4: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 1$, $k = 2$) – Assume Unknown Sensitivity and Specificity



A.2 Additional Results using Weibull distribution with parameter $h = 2$ and $k = 2$

Table A.5: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2115/0.0115	0.2117/0.0117	0.2111/0.0111	0.2121/0.0121	0.2128/0.0128
	SD		0.2697	0.2697	0.2698	0.2694	0.2649
	MSE		0.0728	0.0729	0.0729	0.0727	0.0703
	Relative Efficiency		**1.0000**	**0.9997**	**0.9989**	**1.0018**	**1.0362**
β_2	Mean/Bias	-0.3	-0.3032/-0.0032	-0.3033/-0.0033	-0.3039/-0.0039	-0.3030/-0.0030	-0.3020/-0.0020
	SD		0.2620	0.2620	0.2614	0.2614	0.2562
	MSE		0.0687	0.0687	0.0683	0.0683	0.0657
	Relative Efficiency		**1.0000**	**1.0002**	**1.0053**	**1.0051**	**1.0458**
β_3	Mean/Bias	0.1	0.1067/0.0067	0.1065/0.0065	0.1067/0.0067	0.1062/0.0062	0.1063/0.0063
	SD		0.2120	0.2120	0.2113	0.2117	0.2063
	MSE		0.0450	0.0450	0.0447	0.0448	0.0426
	Relative Efficiency		**1.0000**	**1.0003**	**1.0072**	**1.0032**	**1.0565**
γ_0	Mean/Bias	0.25	0.2476/-0.0024	0.2473/-0.0027	0.2474/-0.0026	0.2473/-0.0027	0.2475/-0.0025
	SD		0.3142	0.3144	0.3131	0.3132	0.3072
	MSE		0.0987	0.0989	0.0981	0.0981	0.0944
	Relative Efficiency		**1.0000**	**0.9985**	**1.0066**	**1.0064**	**1.0457**
γ_1	Mean/Bias	-0.1	-0.1048/-0.0048	-0.1045/-0.0045	-0.1045/-0.0045	-0.1053/-0.0053	-0.1079/-0.0079
	SD		0.3607	0.3613	0.3602	0.3584	0.3524
	MSE		0.1301	0.1306	0.1298	0.1285	0.1242
	Relative Efficiency		**1.0000**	**0.9963**	**1.0026**	**1.0128**	**1.0476**
γ_2	Mean/Bias	0.5	0.5171/0.0171	0.5168/0.0168	0.5170/0.0170	0.5141/0.0141	0.5107/0.0107
	SD		0.3823	0.3823	0.3802	0.3774	0.3685
	MSE		0.1464	0.1464	0.1448	0.1426	0.1359
	Relative Efficiency		**1.0000**	**0.9997**	**1.0109**	**1.0258**	**1.0758**
γ_3	Mean/Bias	-0.1	-0.0917/0.0083	-0.0921/0.0079	-0.0920/0.0080	-0.0920/0.0080	-0.0923/0.0077
	SD		0.3261	0.3259	0.3240	0.3219	0.3156
	MSE		0.1064	0.1063	0.1050	0.1037	0.0997
	Relative Efficiency		**1.0000**	**1.0012**	**1.0130**	**1.0266**	**1.0675**

Figure A.5: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity

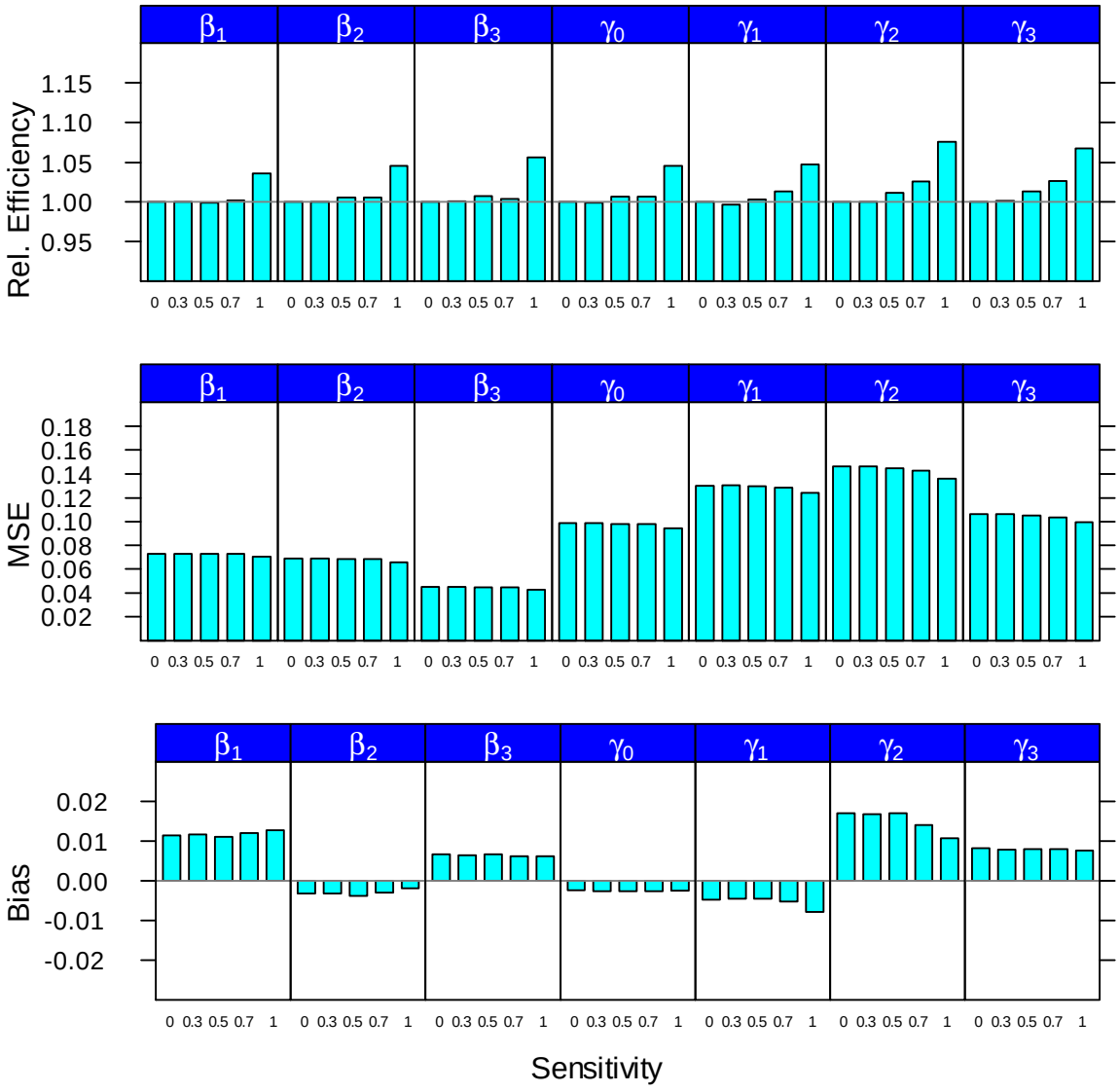


Table A.6: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2115/0.0115	0.2115/0.0115	0.2116/0.0116	0.2125/0.0125	0.2115/0.0115
	SD		0.2697	0.2696	0.2702	0.2709	0.2667
	MSE		0.0728	0.0728	0.0732	0.0735	0.0713
	Relative Efficiency		**1.0000**	**1.0005**	**0.9958**	**0.9909**	**1.0221**
β_2	Mean/Bias	-0.3	-0.3032/-0.0032	-0.3022/-0.0022	-0.3026/-0.0026	-0.3040/-0.0040	-0.3035/-0.0035
	SD		0.2620	0.2618	0.2623	0.2625	0.2564
	MSE		0.0687	0.0685	0.0688	0.0689	0.0658
	Relative Efficiency		**1.0000**	**1.0021**	**0.9983**	**0.9967**	**1.0441**
β_3	Mean/Bias	0.1	0.1067/0.0067	0.1069/0.0069	0.1058/0.0058	0.1077/0.0077	0.1086/0.0086
	SD		0.2120	0.2123	0.2123	0.2128	0.2058
	MSE		0.0450	0.0451	0.0451	0.0454	0.0424
	Relative Efficiency		**1.0000**	**0.9976**	**0.9970**	**0.9923**	**1.0608**
γ_0	Mean/Bias	0.25	0.2476/-0.0024	0.2515/0.0015	0.2522/0.0022	0.2529/0.0029	0.2560/0.0060
	SD		0.3142	0.3149	0.3139	0.3160	0.3109
	MSE		0.0987	0.0991	0.0985	0.0999	0.0967
	Relative Efficiency		**1.0000**	**0.9956**	**1.0017**	**0.9885**	**1.0208**
γ_1	Mean/Bias	-0.1	-0.1048/-0.0048	-0.1063/-0.0063	-0.1067/-0.0067	-0.1068/-0.0068	-0.1090/-0.0090
	SD		0.3607	0.3634	0.3626	0.3599	0.3544
	MSE		0.1301	0.1321	0.1315	0.1296	0.1256
	Relative Efficiency		**1.0000**	**0.9853**	**0.9894**	**1.0044**	**1.0360**
γ_2	Mean/Bias	0.5	0.5171/0.0171	0.5199/0.0199	0.5180/0.0180	0.5170/0.0170	0.5179/0.0179
	SD		0.3823	0.3859	0.3831	0.3810	0.3725
	MSE		0.1464	0.1493	0.1471	0.1454	0.1391
	Relative Efficiency		**1.0000**	**0.9813**	**0.9954**	**1.0067**	**1.0531**
γ_3	Mean/Bias	-0.1	-0.0917/0.0083	-0.0966/0.0034	-0.0917/0.0083	-0.0938/0.0062	-0.0947/0.0053
	SD		0.3261	0.3255	0.3273	0.3248	0.3179
	MSE		0.1064	0.1060	0.1072	0.1055	0.1011
	Relative Efficiency		**1.0000**	**1.0037**	**0.9924**	**1.0082**	**1.0525**

Figure A.6: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity

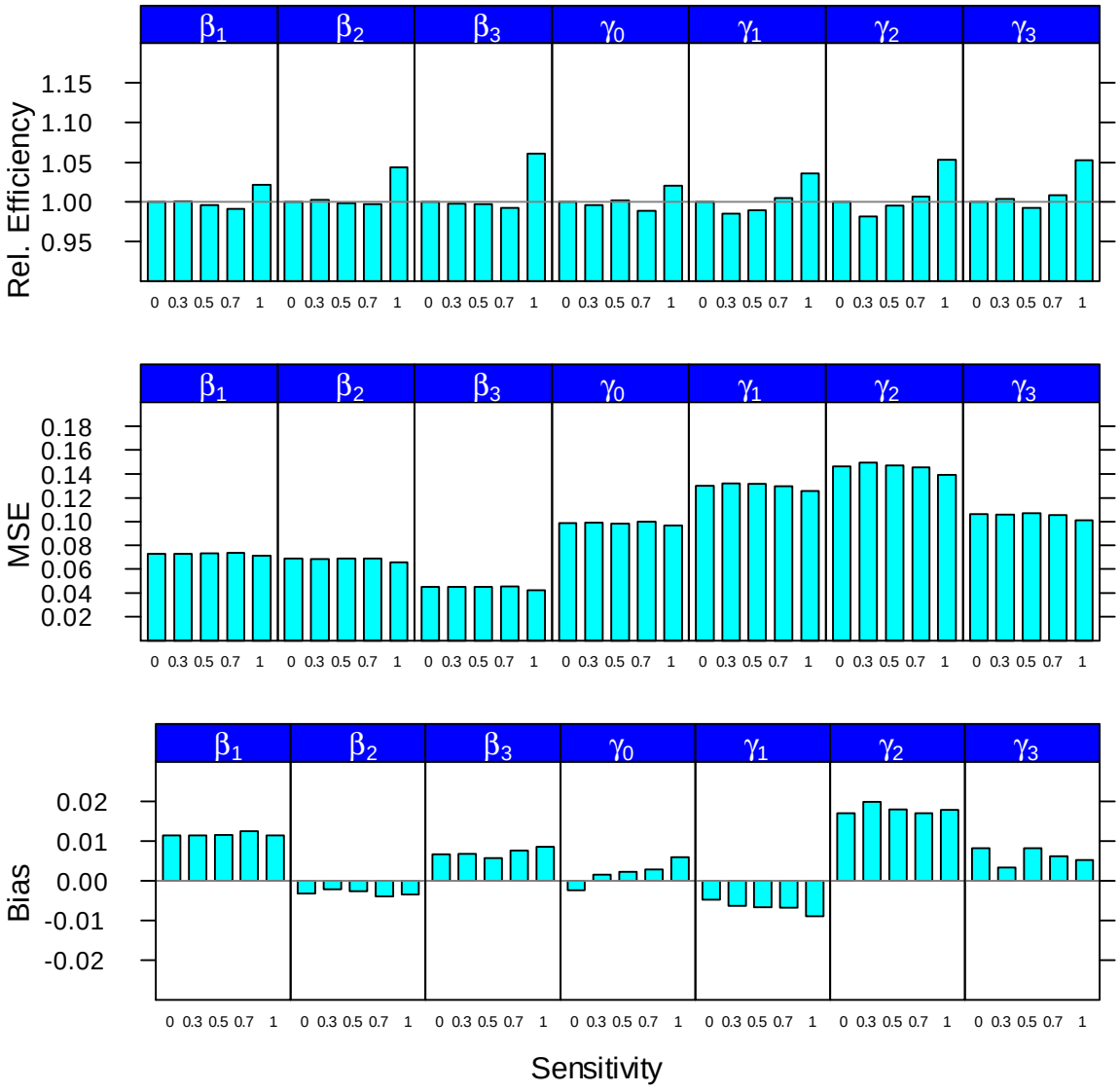


Table A.7: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2115/0.0115	0.2118/0.0118	0.2112/0.0112	0.2125/0.0125	0.2134/0.0134
	SD		0.2697	0.2696	0.2698	0.2692	0.2640
	MSE		0.0728	0.0728	0.0729	0.0726	0.0699
	Relative Efficiency		**1.0000**	**1.0004**	**0.9991**	**1.0032**	**1.0431**
β_2	Mean/Bias	-0.3	-0.3032/-0.0032	-0.3034/-0.0034	-0.3042/-0.0042	-0.3030/-0.0030	-0.3016/-0.0016
	SD		0.2620	0.2618	0.2611	0.2611	0.2555
	MSE		0.0687	0.0686	0.0682	0.0682	0.0653
	Relative Efficiency		**1.0000**	**1.0017**	**1.0072**	**1.0069**	**1.0516**
β_3	Mean/Bias	0.1	0.1067/0.0067	0.1066/0.0066	0.1067/0.0067	0.1060/0.0060	0.1065/0.0065
	SD		0.2120	0.2121	0.2108	0.2114	0.2057
	MSE		0.0450	0.0450	0.0445	0.0447	0.0423
	Relative Efficiency		**1.0000**	**0.9989**	**1.0112**	**1.0060**	**1.0627**
γ_0	Mean/Bias	0.25	0.2476/-0.0024	0.2476/-0.0024	0.2476/-0.0024	0.2477/-0.0023	0.2485/-0.0015
	SD		0.3142	0.3142	0.3120	0.3120	0.3054
	MSE		0.0987	0.0987	0.0973	0.0973	0.0933
	Relative Efficiency		**1.0000**	**1.0001**	**1.0141**	**1.0140**	**1.0582**
γ_1	Mean/Bias	-0.1	-0.1048/-0.0048	-0.1049/-0.0049	-0.1054/-0.0054	-0.1067/-0.0067	-0.1102/-0.0102
	SD		0.3607	0.3612	0.3588	0.3563	0.3492
	MSE		0.1301	0.1305	0.1288	0.1270	0.1220
	Relative Efficiency		**1.0000**	**0.9973**	**1.0102**	**1.0250**	**1.0671**
γ_2	Mean/Bias	0.5	0.5171/0.0171	0.5161/0.0161	0.5175/0.0175	0.5140/0.0140	0.5105/0.0105
	SD		0.3823	0.3816	0.3783	0.3745	0.3654
	MSE		0.1464	0.1459	0.1434	0.1405	0.1336
	Relative Efficiency		**1.0000**	**1.0032**	**1.0210**	**1.0416**	**1.0943**
γ_3	Mean/Bias	-0.1	-0.0917/0.0083	-0.0928/0.0072	-0.0917/0.0083	-0.0917/0.0083	-0.0920/0.0080
	SD		0.3261	0.3251	0.3225	0.3200	0.3135
	MSE		0.1064	0.1058	0.1041	0.1025	0.0983
	Relative Efficiency		**1.0000**	**1.0060**	**1.0226**	**1.0382**	**1.0821**

Figure A.7: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Known Sensitivity and Specificity

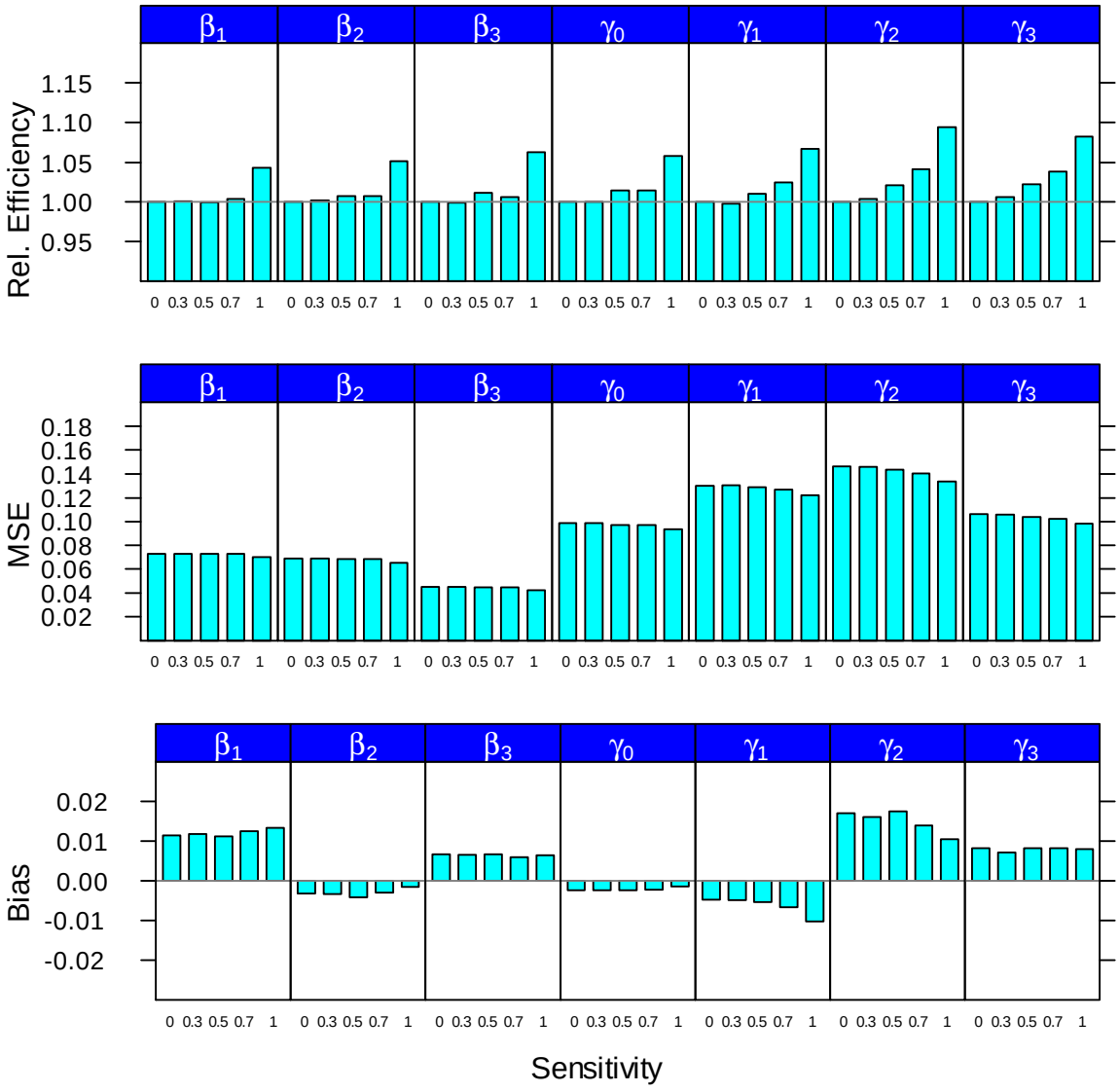
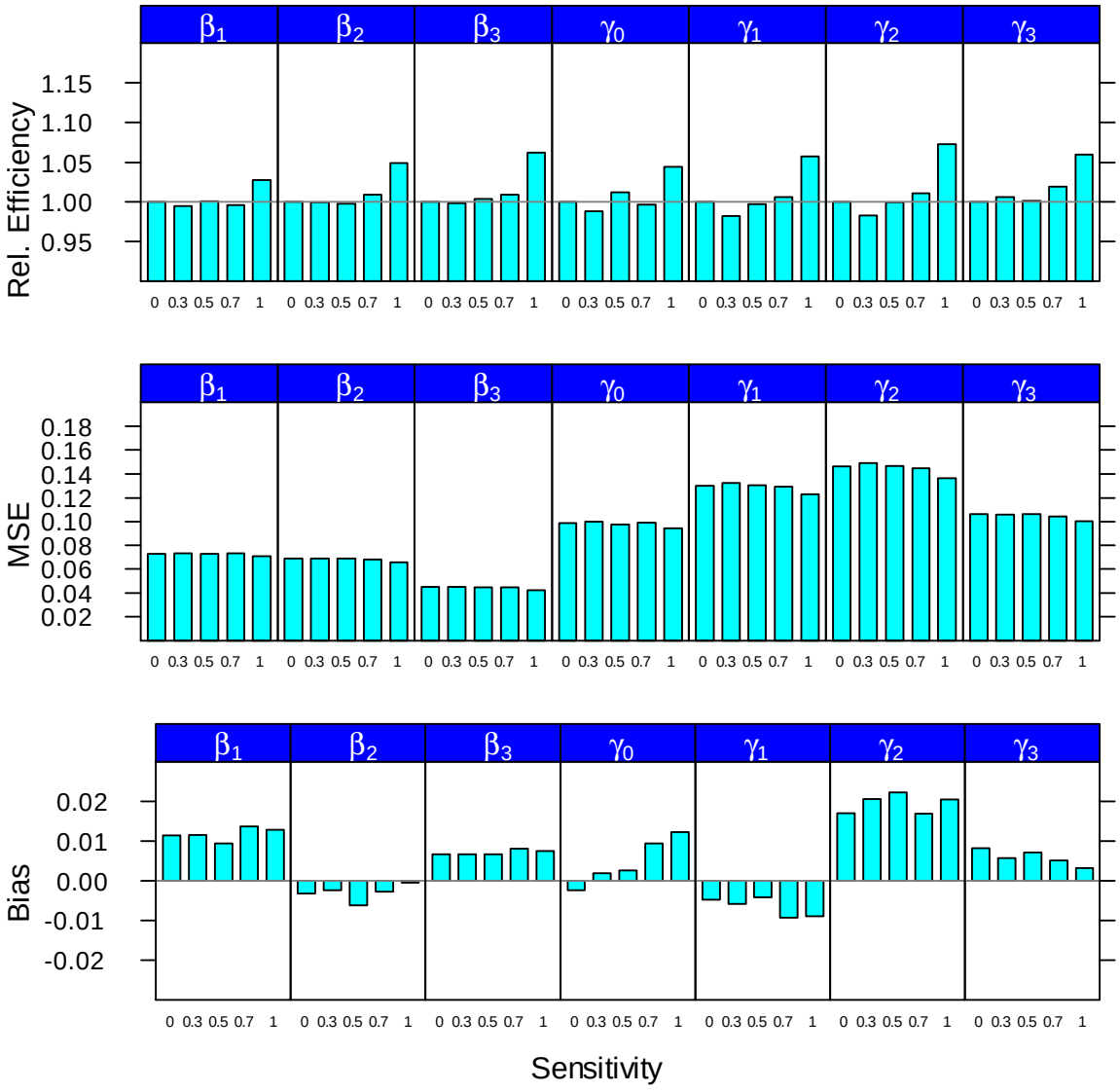


Table A.8: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2115/0.0115	0.2116/0.0116	0.2095/0.0095	0.2137/0.0137	0.2129/0.0129
	SD		0.2697	0.2704	0.2696	0.2702	0.2660
	MSE		0.0728	0.0732	0.0728	0.0732	0.0709
	Relative Efficiency		**1.0000**	**0.9948**	**1.0004**	**0.9959**	**1.0276**
β_2	Mean/Bias	-0.3	-0.3032/-0.0032	-0.3024/-0.0024	-0.3062/-0.0062	-0.3028/-0.0028	-0.3005/-0.0005
	SD		0.2620	0.2621	0.2623	0.2609	0.2558
	MSE		0.0687	0.0687	0.0688	0.0681	0.0655
	Relative Efficiency		**1.0000**	**0.9993**	**0.9978**	**1.0090**	**1.0491**
β_3	Mean/Bias	0.1	0.1067/0.0067	0.1067/0.0067	0.1067/0.0067	0.1081/0.0081	0.1075/0.0075
	SD		0.2120	0.2122	0.2116	0.2111	0.2057
	MSE		0.0450	0.0451	0.0448	0.0446	0.0424
	Relative Efficiency		**1.0000**	**0.9980**	**1.0036**	**1.0087**	**1.0624**
γ_0	Mean/Bias	0.25	0.2476/-0.0024	0.2518/0.0018	0.2526/0.0026	0.2594/0.0094	0.2623/0.0123
	SD		0.3142	0.3160	0.3123	0.3147	0.3074
	MSE		0.0987	0.0999	0.0975	0.0992	0.0946
	Relative Efficiency		**1.0000**	**0.9882**	**1.0120**	**0.9963**	**1.0446**
γ_1	Mean/Bias	-0.1	-0.1048/-0.0048	-0.1058/-0.0058	-0.1042/-0.0042	-0.1093/-0.0093	-0.1090/-0.0090
	SD		0.3607	0.3640	0.3613	0.3597	0.3507
	MSE		0.1301	0.1325	0.1305	0.1294	0.1231
	Relative Efficiency		**1.0000**	**0.9819**	**0.9967**	**1.0057**	**1.0577**
γ_2	Mean/Bias	0.5	0.5171/0.0171	0.5206/0.0206	0.5223/0.0223	0.5169/0.0169	0.5205/0.0205
	SD		0.3823	0.3856	0.3823	0.3802	0.3691
	MSE		0.1464	0.1491	0.1467	0.1448	0.1366
	Relative Efficiency		**1.0000**	**0.9826**	**0.9996**	**1.0108**	**1.0728**
γ_3	Mean/Bias	-0.1	-0.0917/0.0083	-0.0942/0.0058	-0.0928/0.0072	-0.0948/0.0052	-0.0967/0.0033
	SD		0.3261	0.3252	0.3259	0.3229	0.3168
	MSE		0.1064	0.1058	0.1063	0.1043	0.1004
	Relative Efficiency		**1.0000**	**1.0058**	**1.0011**	**1.0197**	**1.0596**

Figure A.8: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = 2$, $k = 2$) – Assume Unknown Sensitivity and Specificity



A.3 Additional Results using Weibull distribution with parameter $h = \frac{2}{3}$ and $k = 3$

Table A.9: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2097/0.0097	0.2095/0.0095	0.2093/0.0093	0.2087/0.0087	0.2105/0.0105
	SD		0.3146	0.3144	0.3132	0.3109	0.2964
	MSE		0.0991	0.0989	0.0982	0.0968	0.0880
	Relative Efficiency		**1.0000**	**1.0017**	**1.0089**	**1.0238**	**1.1264**
β_2	Mean/Bias	-0.3	-0.3131/-0.0131	-0.3131/-0.0131	-0.3139/-0.0139	-0.3125/-0.0125	-0.3096/-0.0096
	SD		0.3028	0.3026	0.3002	0.2986	0.2887
	MSE		0.0918	0.0917	0.0903	0.0893	0.0834
	Relative Efficiency		**1.0000**	**1.0012**	**1.0176**	**1.0284**	**1.1002**
β_3	Mean/Bias	0.1	0.1114/0.0114	0.1107/0.0107	0.1118/0.0118	0.1100/0.0100	0.1090/0.0090
	SD		0.2513	0.2507	0.2497	0.2480	0.2330
	MSE		0.0633	0.0630	0.0625	0.0616	0.0544
	Relative Efficiency		**1.0000**	**1.0048**	**1.0125**	**1.0268**	**1.1632**
γ_0	Mean/Bias	0.25	0.2649/0.0149	0.2654/0.0154	0.2658/0.0158	0.2616/0.0116	0.2591/0.0091
	SD		0.3553	0.3532	0.3454	0.3371	0.3100
	MSE		0.1265	0.1250	0.1196	0.1138	0.0962
	Relative Efficiency		**1.0000**	**1.0123**	**1.0582**	**1.1111**	**1.3136**
γ_1	Mean/Bias	-0.1	-0.0984/0.0016	-0.0993/0.0007	-0.0960/0.0040	-0.0948/0.0052	-0.0955/0.0045
	SD		0.4342	0.4314	0.4247	0.4128	0.3816
	MSE		0.1885	0.1861	0.1804	0.1704	0.1456
	Relative Efficiency		**1.0000**	**1.0128**	**1.0451**	**1.1063**	**1.2949**
γ_2	Mean/Bias	0.5	0.4998/-0.0002	0.4996/-0.0004	0.5022/0.0022	0.4991/-0.0009	0.4995/-0.0005
	SD		0.4503	0.4500	0.4437	0.4299	0.4016
	MSE		0.2027	0.2025	0.1969	0.1848	0.1613
	Relative Efficiency		**1.0000**	**1.0012**	**1.0297**	**1.0969**	**1.2572**
γ_3	Mean/Bias	-0.1	-0.1152/-0.0152	-0.1151/-0.0151	-0.1175/-0.0175	-0.1127/-0.0127	-0.1102/-0.0102
	SD		0.3486	0.3479	0.3414	0.3355	0.3113
	MSE		0.1218	0.1212	0.1169	0.1127	0.0970
	Relative Efficiency		**1.0000**	**1.0044**	**1.0430**	**1.0801**	**1.2544**

Figure A.9: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity

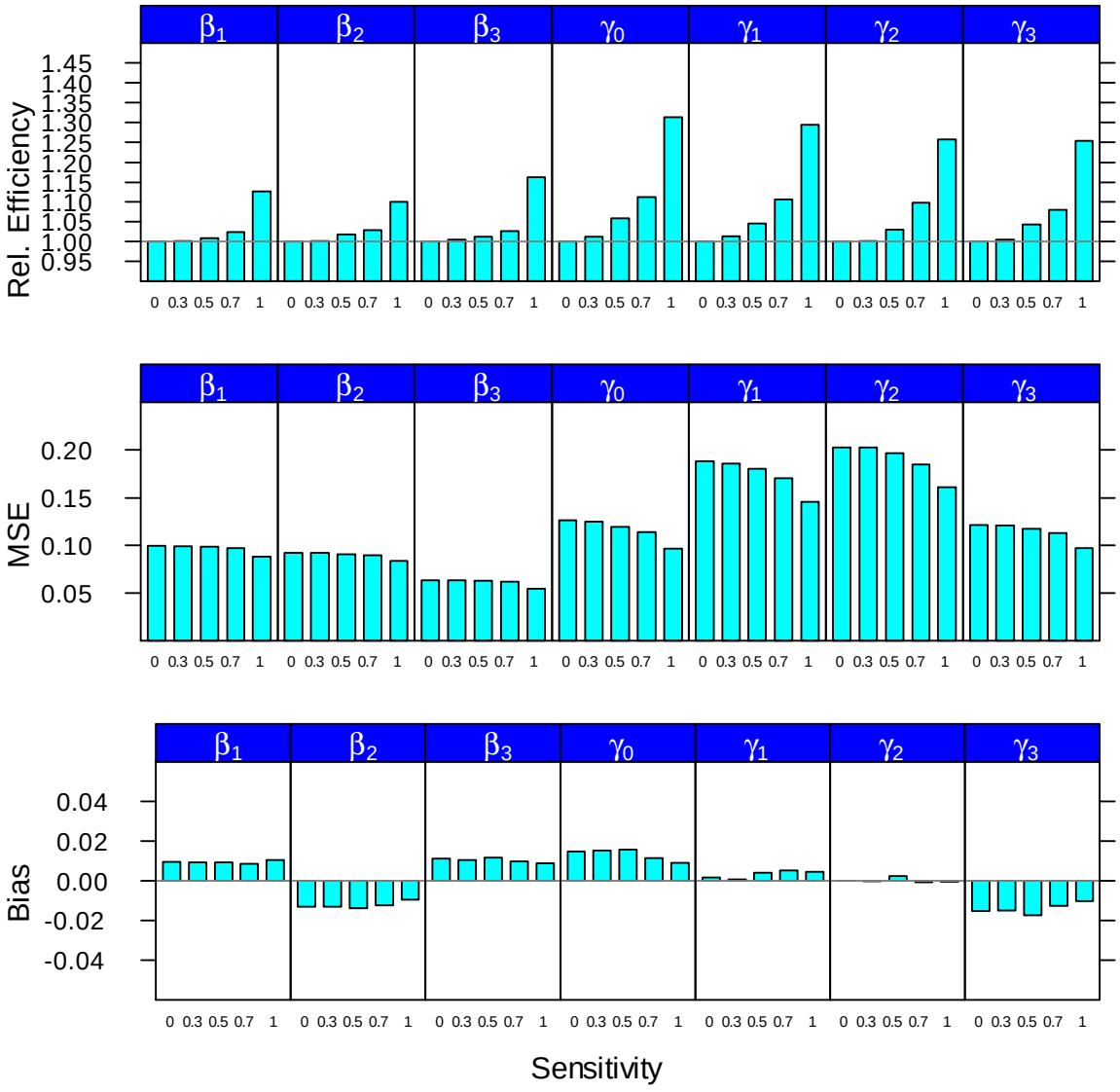


Table A.10: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2097/0.0097	0.2087/0.0087	0.2115/0.0115	0.2111/0.0111	0.2109/0.0109
	SD		0.3146	0.3157	0.3124	0.3112	0.2972
	MSE		0.0991	0.0998	0.0977	0.0969	0.0884
	Relative Efficiency		**1.0000**	**0.9930**	**1.0141**	**1.0223**	**1.1208**
β_2	Mean/Bias	-0.3	-0.3131/-0.0131	-0.3136/-0.0136	-0.3136/-0.0136	-0.3127/-0.0127	-0.3107/-0.0107
	SD		0.3028	0.3038	0.2989	0.2998	0.2880
	MSE		0.0918	0.0925	0.0895	0.0901	0.0831
	Relative Efficiency		**1.0000**	**0.9934**	**1.0265**	**1.0198**	**1.1052**
β_3	Mean/Bias	0.1	0.1114/0.0114	0.1129/0.0129	0.1139/0.0139	0.1109/0.0109	0.1095/0.0095
	SD		0.2513	0.2514	0.2517	0.2484	0.2334
	MSE		0.0633	0.0634	0.0636	0.0618	0.0546
	Relative Efficiency		**1.0000**	**0.9989**	**0.9966**	**1.0235**	**1.1589**
γ_0	Mean/Bias	0.25	0.2649/0.0149	0.2673/0.0173	0.2714/0.0214	0.2683/0.0183	0.2721/0.0221
	SD		0.3553	0.3592	0.3530	0.3439	0.3246
	MSE		0.1265	0.1293	0.1251	0.1186	0.1059
	Relative Efficiency		**1.0000**	**0.9787**	**1.0130**	**1.0673**	**1.1980**
γ_1	Mean/Bias	-0.1	-0.0984/0.0016	-0.0983/0.0017	-0.1022/-0.0022	-0.0933/0.0067	-0.0996/0.0004
	SD		0.4342	0.4368	0.4302	0.4202	0.3852
	MSE		0.1885	0.1908	0.1850	0.1766	0.1484
	Relative Efficiency		**1.0000**	**0.9879**	**1.0188**	**1.0675**	**1.2704**
γ_2	Mean/Bias	0.5	0.4998/-0.0002	0.5013/0.0013	0.5064/0.0064	0.5103/0.0103	0.5081/0.0081
	SD		0.4503	0.4547	0.4485	0.4407	0.4101
	MSE		0.2027	0.2068	0.2012	0.1944	0.1682
	Relative Efficiency		**1.0000**	**0.9804**	**1.0080**	**1.0437**	**1.2055**
γ_3	Mean/Bias	-0.1	-0.1152/-0.0152	-0.1159/-0.0159	-0.1201/-0.0201	-0.1184/-0.0184	-0.1170/-0.0170
	SD		0.3486	0.3521	0.3454	0.3411	0.3189
	MSE		0.1218	0.1242	0.1197	0.1167	0.1020
	Relative Efficiency		**1.0000**	**0.9805**	**1.0189**	**1.0448**	**1.1954**

Figure A.10: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity

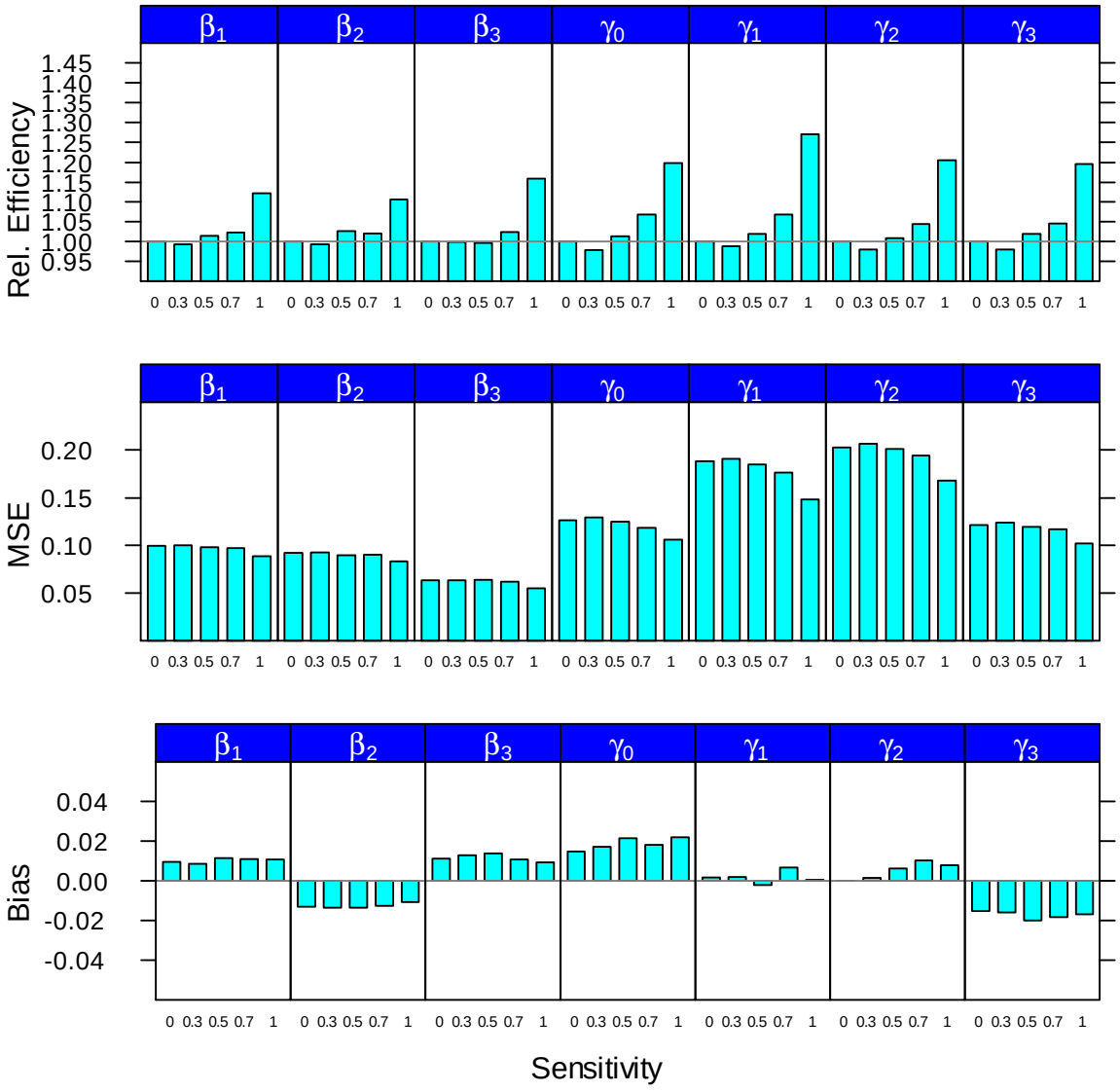


Table A.11: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2097/0.0097	0.2096/0.0096	0.2093/0.0093	0.2086/0.0086	0.2108/0.0108
	SD		0.3146	0.3139	0.3125	0.3100	0.2949
	MSE		0.0991	0.0986	0.0978	0.0962	0.0871
	Relative Efficiency		**1.0000**	**1.0048**	**1.0135**	**1.0299**	**1.1386**
β_2	Mean/Bias	-0.3	-0.3131/-0.0131	-0.3132/-0.0132	-0.3142/-0.0142	-0.3123/-0.0123	-0.3102/-0.0102
	SD		0.3028	0.3018	0.2989	0.2972	0.2876
	MSE		0.0918	0.0913	0.0895	0.0885	0.0828
	Relative Efficiency		**1.0000**	**1.0063**	**1.0263**	**1.0381**	**1.1084**
β_3	Mean/Bias	0.1	0.1114/0.0114	0.1101/0.0101	0.1116/0.0116	0.1094/0.0094	0.1080/0.0080
	SD		0.2513	0.2501	0.2489	0.2470	0.2323
	MSE		0.0633	0.0626	0.0621	0.0611	0.0540
	Relative Efficiency		**1.0000**	**1.0099**	**1.0191**	**1.0351**	**1.1702**
γ_0	Mean/Bias	0.25	0.2649/0.0149	0.2634/0.0134	0.2636/0.0136	0.2576/0.0076	0.2543/0.0043
	SD		0.3553	0.3482	0.3386	0.3296	0.3028
	MSE		0.1265	0.1214	0.1148	0.1087	0.0917
	Relative Efficiency		**1.0000**	**1.0413**	**1.1012**	**1.1620**	**1.3774**
γ_1	Mean/Bias	-0.1	-0.0984/0.0016	-0.0994/0.0006	-0.0953/0.0047	-0.0942/0.0058	-0.0956/0.0044
	SD		0.4342	0.4265	0.4186	0.4059	0.3750
	MSE		0.1885	0.1819	0.1753	0.1648	0.1406
	Relative Efficiency		**1.0000**	**1.0362**	**1.0758**	**1.1441**	**1.3407**
γ_2	Mean/Bias	0.5	0.4998/-0.0002	0.5003/0.0003	0.5029/0.0029	0.4990/-0.0010	0.4997/-0.0003
	SD		0.4503	0.4454	0.4353	0.4178	0.3906
	MSE		0.2027	0.1984	0.1895	0.1746	0.1525
	Relative Efficiency		**1.0000**	**1.0220**	**1.0702**	**1.1612**	**1.3292**
γ_3	Mean/Bias	-0.1	-0.1152/-0.0152	-0.1136/-0.0136	-0.1162/-0.0162	-0.1093/-0.0093	-0.1061/-0.0061
	SD		0.3486	0.3442	0.3343	0.3275	0.3029
	MSE		0.1218	0.1186	0.1120	0.1073	0.0918
	Relative Efficiency		**1.0000**	**1.0262**	**1.0876**	**1.1334**	**1.3245**

Figure A.11: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Known Sensitivity and Specificity

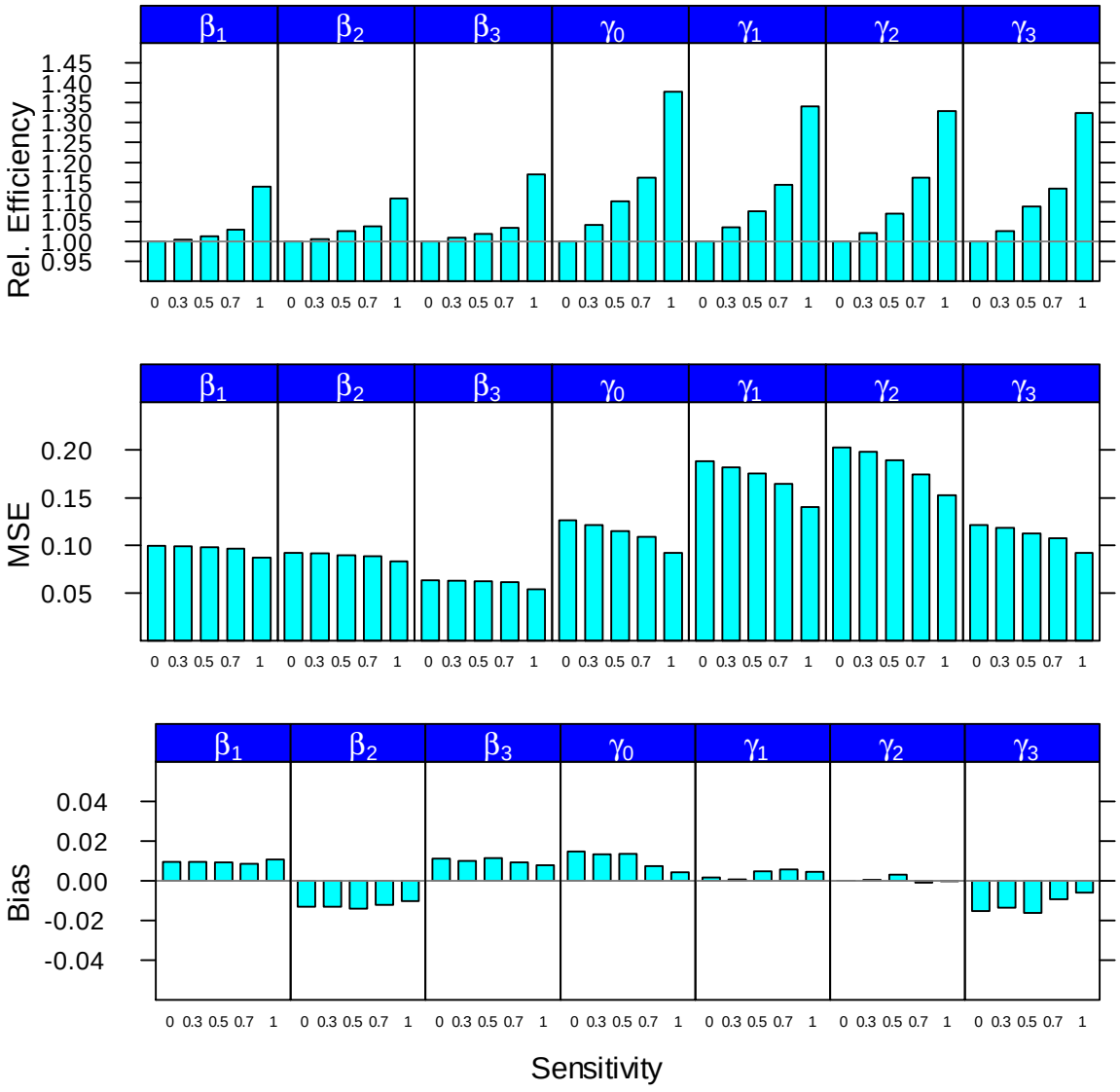
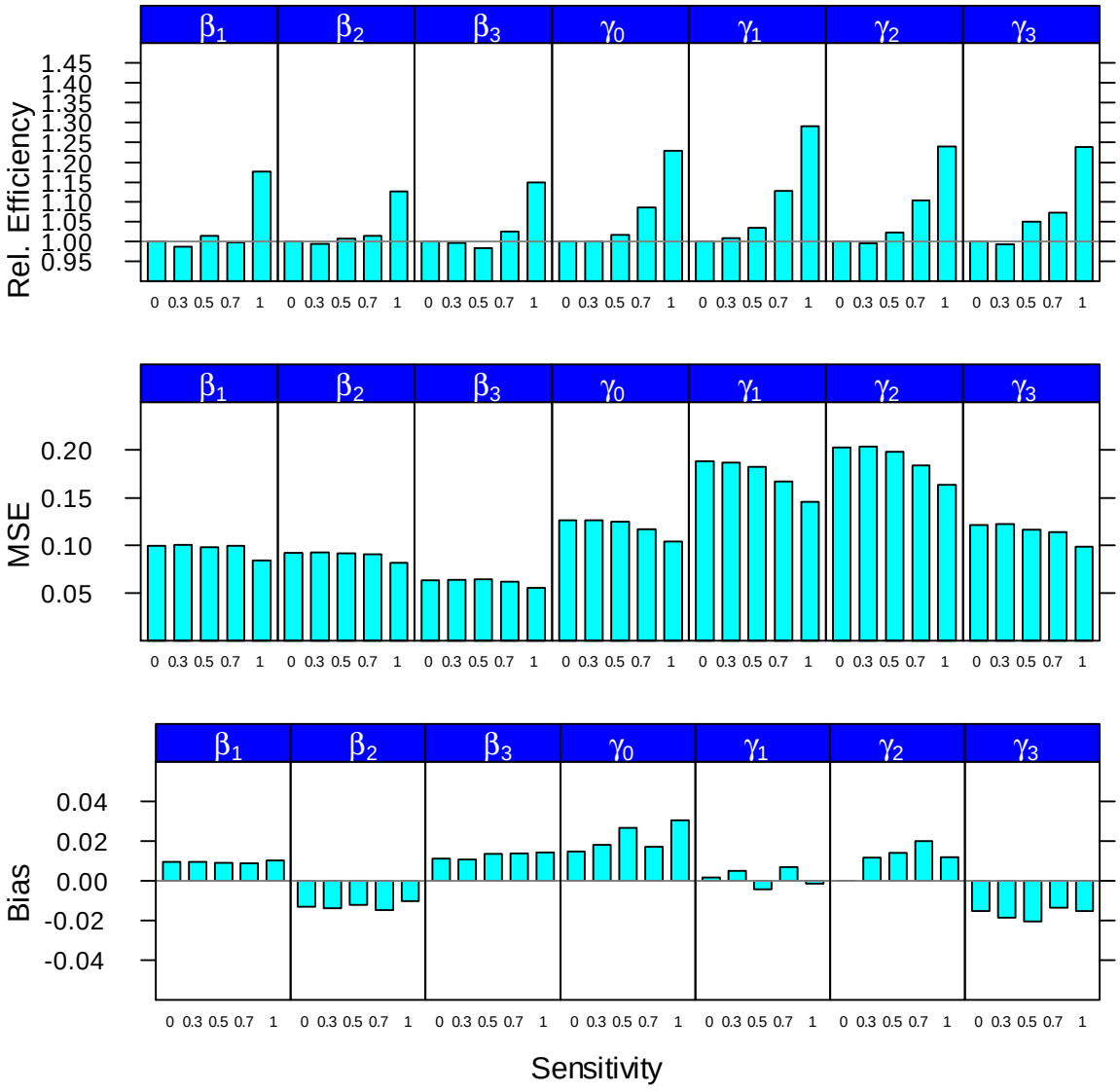


Table A.12: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2097/0.0097	0.2097/0.0097	0.2091/0.0091	0.2090/0.0090	0.2103/0.0103
	SD		0.3146	0.3168	0.3124	0.3150	0.2899
	MSE		0.0991	0.1004	0.0977	0.0993	0.0842
	Relative Efficiency		**1.0000**	**0.9865**	**1.0141**	**0.9973**	**1.1775**
β_2	Mean/Bias	-0.3	-0.3131/-0.0131	-0.3138/-0.0138	-0.3122/-0.0122	-0.3147/-0.0147	-0.3102/-0.0102
	SD		0.3028	0.3037	0.3018	0.3007	0.2854
	MSE		0.0918	0.0925	0.0912	0.0906	0.0815
	Relative Efficiency		**1.0000**	**0.9937**	**1.0067**	**1.0139**	**1.1258**
β_3	Mean/Bias	0.1	0.1114/0.0114	0.1108/0.0108	0.1137/0.0137	0.1139/0.0139	0.1143/0.0143
	SD		0.2513	0.2518	0.2533	0.2482	0.2344
	MSE		0.0633	0.0635	0.0644	0.0618	0.0552
	Relative Efficiency		**1.0000**	**0.9961**	**0.9839**	**1.0251**	**1.1491**
γ_0	Mean/Bias	0.25	0.2649/0.0149	0.2681/0.0181	0.2768/0.0268	0.2673/0.0173	0.2805/0.0305
	SD		0.3553	0.3554	0.3525	0.3410	0.3204
	MSE		0.1265	0.1266	0.1250	0.1166	0.1036
	Relative Efficiency		**1.0000**	**0.9997**	**1.0162**	**1.0857**	**1.2297**
γ_1	Mean/Bias	-0.1	-0.0984/0.0016	-0.0950/0.0050	-0.1044/-0.0044	-0.0929/0.0071	-0.1015/-0.0015
	SD		0.4342	0.4324	0.4269	0.4089	0.3821
	MSE		0.1885	0.1870	0.1823	0.1672	0.1460
	Relative Efficiency		**1.0000**	**1.0084**	**1.0343**	**1.1277**	**1.2909**
γ_2	Mean/Bias	0.5	0.4998/-0.0002	0.5119/0.0119	0.5141/0.0141	0.5200/0.0200	0.5121/0.0121
	SD		0.4503	0.4513	0.4452	0.4286	0.4045
	MSE		0.2027	0.2038	0.1984	0.1841	0.1637
	Relative Efficiency		**1.0000**	**0.9954**	**1.0228**	**1.1036**	**1.2393**
γ_3	Mean/Bias	-0.1	-0.1152/-0.0152	-0.1185/-0.0185	-0.1205/-0.0205	-0.1136/-0.0136	-0.1152/-0.0152
	SD		0.3486	0.3499	0.3403	0.3366	0.3133
	MSE		0.1218	0.1228	0.1162	0.1135	0.0984
	Relative Efficiency		**1.0000**	**0.9929**	**1.0498**	**1.0726**	**1.2386**

Figure A.12: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{2}{3}$, $k = 3$) – Assume Unknown Sensitivity and Specificity



A.4 Additional Results using Weibull distribution with parameter $h = \frac{1}{3}$ and $k = 4$

Table A.13: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2037/0.0037	0.2026/0.0026	0.2059/0.0059	0.2066/0.0066	0.1998/-0.0002
	SD		0.4272	0.4271	0.4216	0.4122	0.3726
	MSE		0.1825	0.1824	0.1778	0.1699	0.1388
	Relative Efficiency		**1.0000**	**1.0005**	**1.0267**	**1.0743**	**1.3148**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3029/-0.0029	-0.3014/-0.0014	-0.3025/-0.0025	-0.3016/-0.0016
	SD		0.4029	0.4010	0.3968	0.3841	0.3472
	MSE		0.1623	0.1608	0.1574	0.1476	0.1206
	Relative Efficiency		**1.0000**	**1.0098**	**1.0311**	**1.1003**	**1.3465**
β_3	Mean/Bias	0.1	0.1213/0.0213	0.1198/0.0198	0.1203/0.0203	0.1192/0.0192	0.1184/0.0184
	SD		0.3080	0.3085	0.3055	0.2985	0.2738
	MSE		0.0953	0.0956	0.0937	0.0895	0.0753
	Relative Efficiency		**1.0000**	**0.9968**	**1.0163**	**1.0645**	**1.2658**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2555/0.0055	0.2563/0.0063	0.2542/0.0042	0.2500/0.0000
	SD		0.4451	0.4363	0.4141	0.3816	0.3179
	MSE		0.1982	0.1904	0.1715	0.1457	0.1010
	Relative Efficiency		**1.0000**	**1.0407**	**1.1554**	**1.3604**	**1.9611**
γ_1	Mean/Bias	-0.1	-0.0937/0.0063	-0.0910/0.0090	-0.0929/0.0071	-0.0955/0.0045	-0.0925/0.0075
	SD		0.5371	0.5328	0.5087	0.4632	0.3919
	MSE		0.2885	0.2840	0.2589	0.2146	0.1536
	Relative Efficiency		**1.0000**	**1.0161**	**1.1145**	**1.3442**	**1.8784**
γ_2	Mean/Bias	0.5	0.4992/-0.0008	0.4967/-0.0033	0.4936/-0.0064	0.4849/-0.0151	0.4817/-0.0183
	SD		0.5823	0.5754	0.5480	0.4952	0.4117
	MSE		0.3391	0.3311	0.3004	0.2454	0.1698
	Relative Efficiency		**1.0000**	**1.0240**	**1.1290**	**1.3827**	**2.0007**
γ_3	Mean/Bias	-0.1	-0.0982/0.0018	-0.0943/0.0057	-0.0926/0.0074	-0.0894/0.0106	-0.0879/0.0121
	SD		0.4735	0.4695	0.4459	0.4126	0.3378
	MSE		0.2242	0.2204	0.1989	0.1704	0.1143
	Relative Efficiency		**1.0000**	**1.0172**	**1.1277**	**1.3167**	**1.9642**

Figure A.13: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity

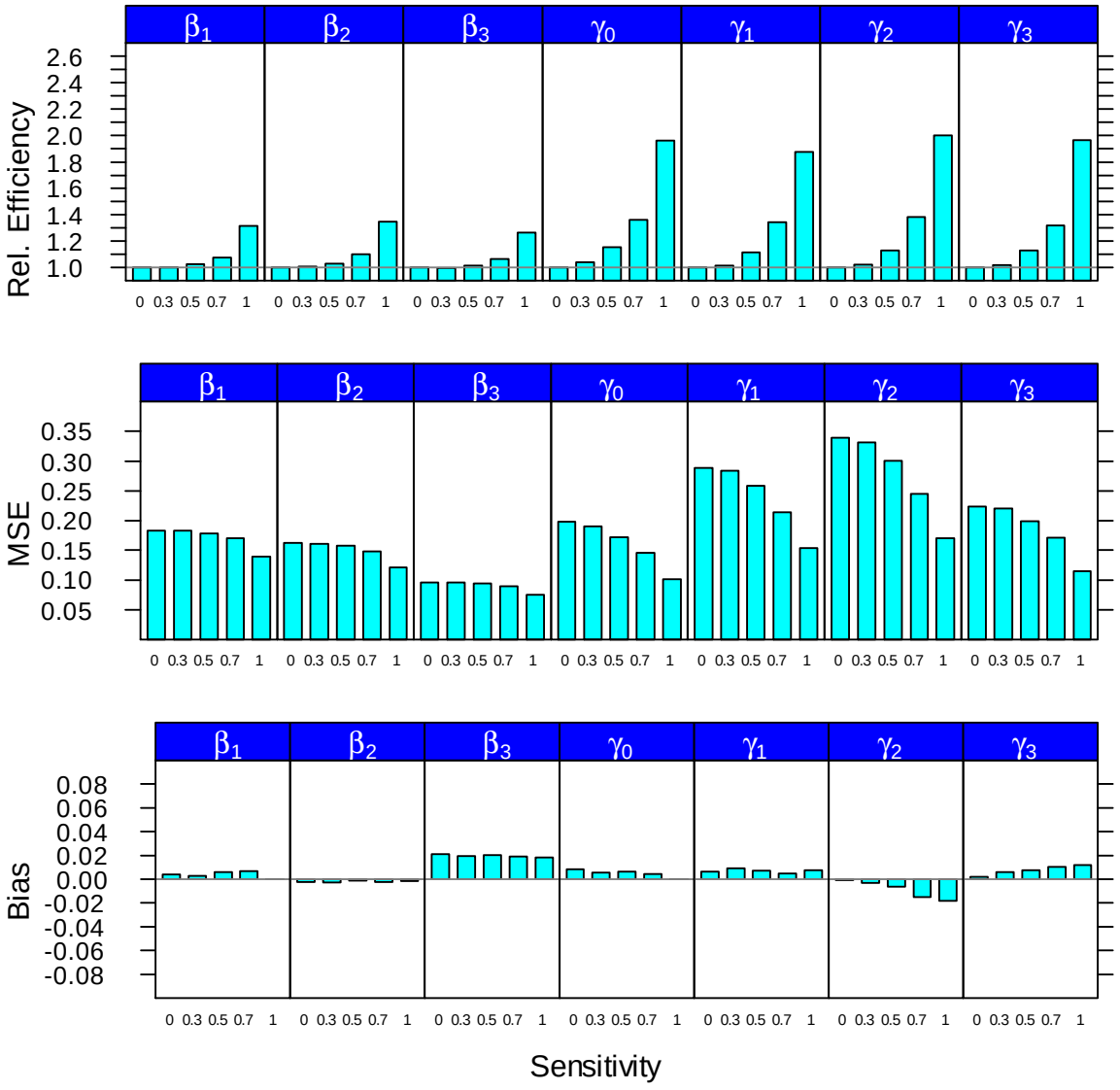


Table A.14: PH Simulation Results for 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2037/0.0037	0.1925/-0.0075	0.2019/0.0019	0.2020/0.0020	0.2028/0.0028
	SD		0.4272	0.4220	0.4209	0.4137	0.3801
	MSE		0.1825	0.1782	0.1772	0.1711	0.1445
	Relative Efficiency		**1.0000**	**1.0245**	**1.0299**	**1.0663**	**1.2630**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3077/-0.0077	-0.3069/-0.0069	-0.3168/-0.0168	-0.3078/-0.0078
	SD		0.4029	0.3958	0.3994	0.3895	0.3509
	MSE		0.1623	0.1567	0.1596	0.1520	0.1232
	Relative Efficiency		**1.0000**	**1.0364**	**1.0175**	**1.0698**	**1.3180**
β_3	Mean/Bias	0.1	0.1213/0.0213	0.1187/0.0187	0.1181/0.0181	0.1170/0.0170	0.1185/0.0185
	SD		0.3080	0.3104	0.3078	0.3007	0.2835
	MSE		0.0953	0.0967	0.0951	0.0907	0.0807
	Relative Efficiency		**1.0000**	**0.9845**	**1.0010**	**1.0489**	**1.1805**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2676/0.0176	0.2632/0.0132	0.2741/0.0241	0.2713/0.0213
	SD		0.4451	0.4486	0.4385	0.4112	0.3524
	MSE		0.1982	0.2016	0.1925	0.1697	0.1247
	Relative Efficiency		**1.0000**	**0.9844**	**1.0302**	**1.1716**	**1.5951**
γ_1	Mean/Bias	-0.1	-0.0937/0.0063	-0.0902/0.0098	-0.0963/0.0037	-0.0886/0.0114	-0.0909/0.0091
	SD		0.5371	0.5463	0.5262	0.4823	0.4007
	MSE		0.2885	0.2985	0.2769	0.2328	0.1606
	Relative Efficiency		**1.0000**	**0.9666**	**1.0417**	**1.2398**	**1.7966**
γ_2	Mean/Bias	0.5	0.4992/-0.0008	0.5006/0.0006	0.5105/0.0105	0.5099/0.0099	0.5023/0.0023
	SD		0.5823	0.6003	0.5768	0.5338	0.4430
	MSE		0.3391	0.3603	0.3329	0.2850	0.1963
	Relative Efficiency		**1.0000**	**0.9409**	**1.0189**	**1.1900**	**1.7273**
γ_3	Mean/Bias	-0.1	-0.0982/0.0018	-0.1019/-0.0019	-0.0956/0.0044	-0.1038/-0.0038	-0.0970/0.0030
	SD		0.4735	0.4805	0.4629	0.4416	0.3493
	MSE		0.2242	0.2309	0.2143	0.1951	0.1220
	Relative Efficiency		**1.0000**	**0.9710**	**1.0462**	**1.1494**	**1.8378**

Figure A.14: Bias and Relative Efficiency for the PH Simulation Results of 80% Specificity ($p_1 = 0.2$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity

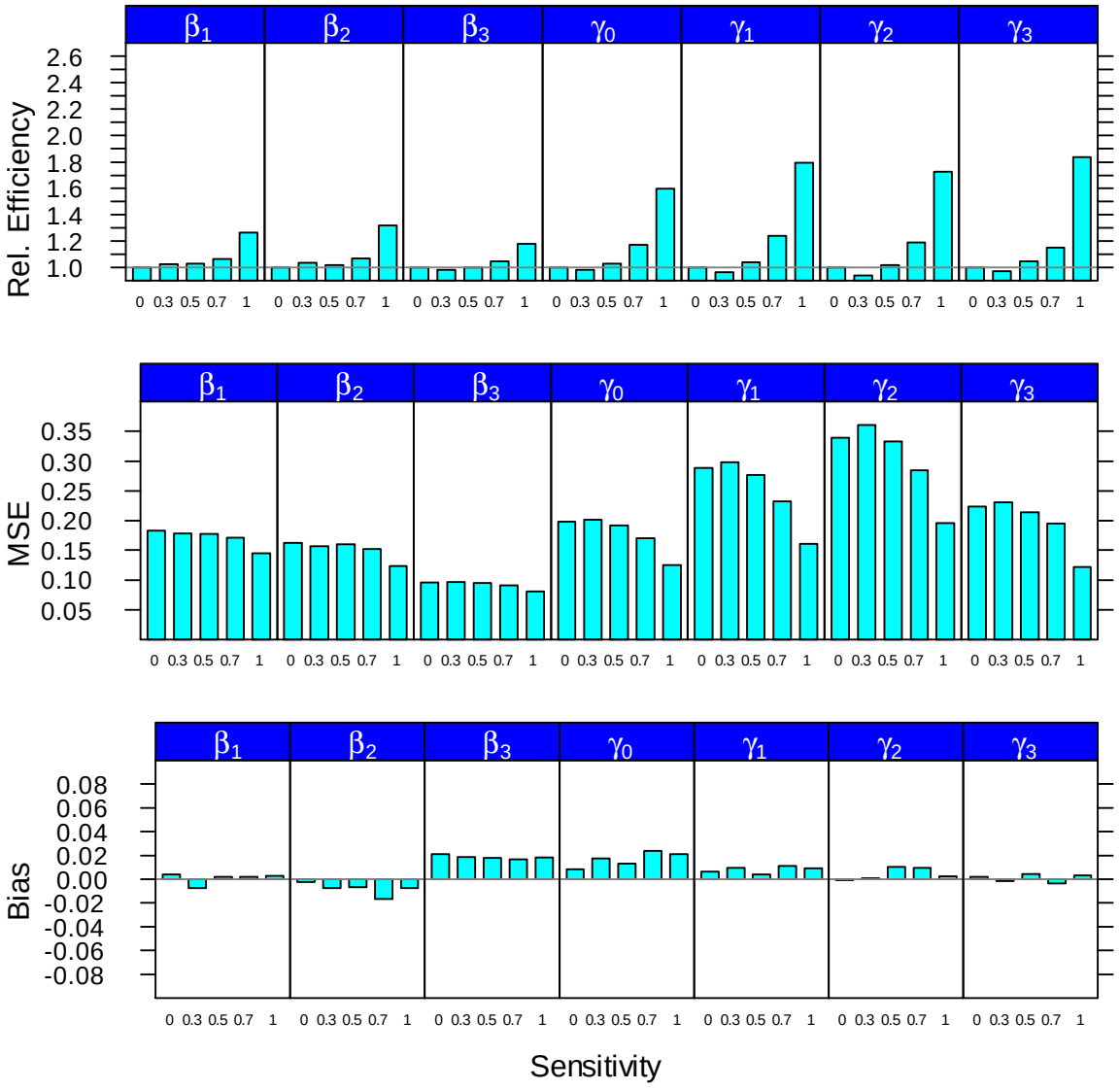


Table A.15: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2037/0.0037	0.2043/0.0043	0.2064/0.0064	0.2073/0.0073	0.1995/-0.0005
	SD		0.4272	0.4269	0.4168	0.4061	0.3664
	MSE		0.1825	0.1822	0.1738	0.1649	0.1343
	Relative Efficiency		**1.0000**	**1.0015**	**1.0503**	**1.1068**	**1.3591**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3004/-0.0004	-0.3010/-0.0010	-0.3020/-0.0020	-0.3015/-0.0015
	SD		0.4029	0.3974	0.3915	0.3767	0.3406
	MSE		0.1623	0.1579	0.1533	0.1419	0.1160
	Relative Efficiency		**1.0000**	**1.0279**	**1.0589**	**1.1438**	**1.3992**
β_3	Mean/Bias	0.1	0.1213/0.0213	0.1194/0.0194	0.1201/0.0201	0.1190/0.0190	0.1182/0.0182
	SD		0.3080	0.3078	0.3035	0.2953	0.2700
	MSE		0.0953	0.0951	0.0925	0.0876	0.0732
	Relative Efficiency		**1.0000**	**1.0013**	**1.0301**	**1.0880**	**1.3016**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2565/0.0065	0.2541/0.0041	0.2531/0.0031	0.2494/-0.0006
	SD		0.4451	0.4249	0.3928	0.3615	0.3033
	MSE		0.1982	0.1806	0.1543	0.1307	0.0920
	Relative Efficiency		**1.0000**	**1.0974**	**1.2843**	**1.5161**	**2.1544**
γ_1	Mean/Bias	-0.1	-0.0937/0.0063	-0.0916/0.0084	-0.0930/0.0070	-0.0970/0.0030	-0.0940/0.0060
	SD		0.5371	0.5204	0.4833	0.4389	0.3743
	MSE		0.2885	0.2709	0.2336	0.1926	0.1402
	Relative Efficiency		**1.0000**	**1.0651**	**1.2349**	**1.4976**	**2.0583**
γ_2	Mean/Bias	0.5	0.4992/-0.0008	0.4920/-0.0080	0.4919/-0.0081	0.4858/-0.0142	0.4844/-0.0156
	SD		0.5823	0.5562	0.5165	0.4651	0.3963
	MSE		0.3391	0.3094	0.2669	0.2166	0.1573
	Relative Efficiency		**1.0000**	**1.0960**	**1.2709**	**1.5671**	**2.1593**
γ_3	Mean/Bias	-0.1	-0.0982/0.0018	-0.0920/0.0080	-0.0918/0.0082	-0.0892/0.0108	-0.0885/0.0115
	SD		0.4735	0.4576	0.4248	0.3895	0.3198
	MSE		0.2242	0.2095	0.1805	0.1519	0.1024
	Relative Efficiency		**1.0000**	**1.0705**	**1.2427**	**1.4774**	**2.1924**

Figure A.15: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity

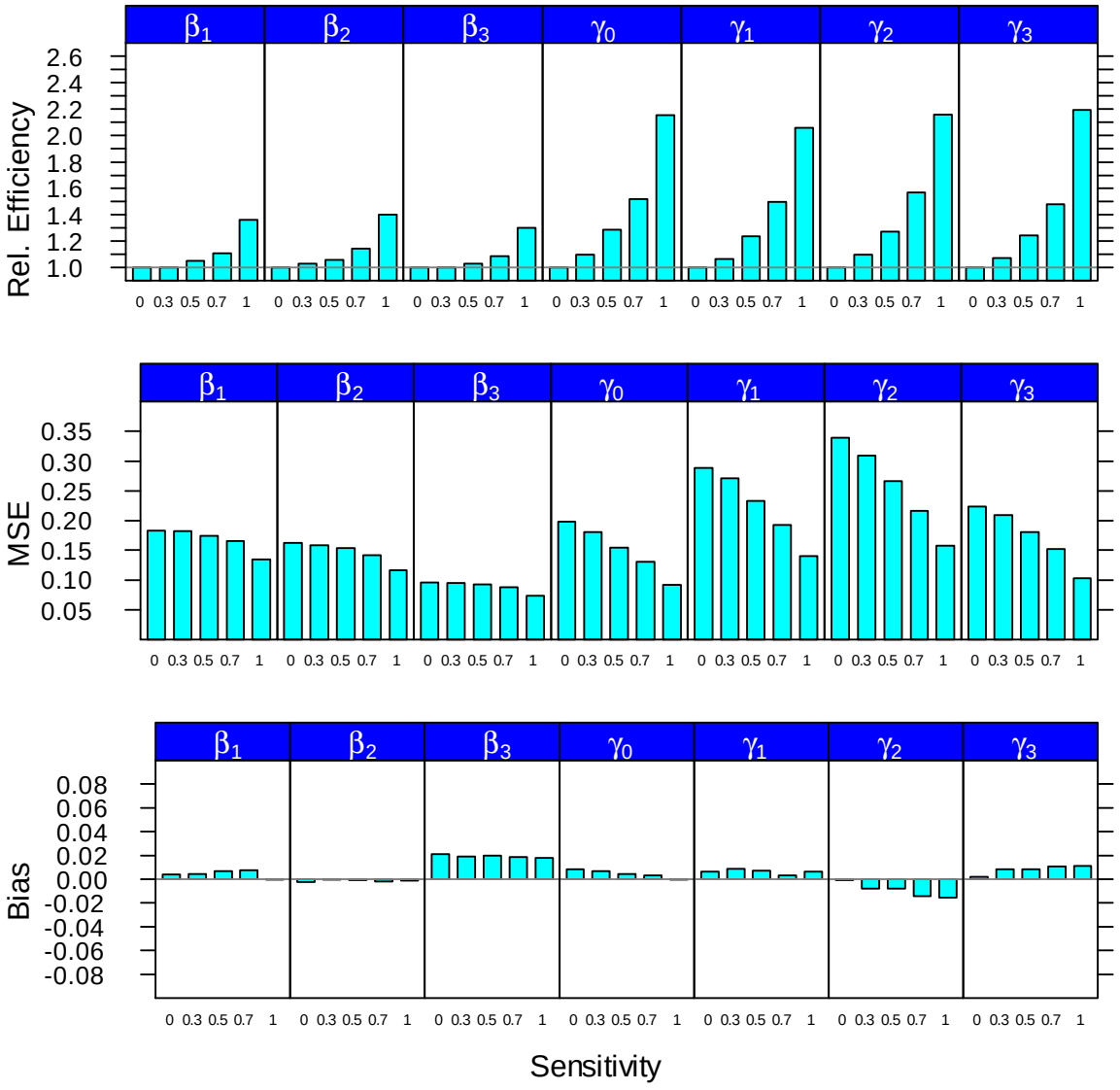


Table A.16: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2037/0.0037	0.1991/-0.0009	0.2036/0.0036	0.2063/0.0063	0.2037/0.0037
	SD		0.4272	0.4250	0.4196	0.4115	0.3655
	MSE		0.1825	0.1806	0.1761	0.1694	0.1336
	Relative Efficiency		**1.0000**	**1.0104**	**1.0363**	**1.0778**	**1.3657**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3028/-0.0028	-0.3067/-0.0067	-0.3134/-0.0134	-0.3110/-0.0110
	SD		0.4029	0.3992	0.3887	0.3827	0.3430
	MSE		0.1623	0.1593	0.1511	0.1466	0.1177
	Relative Efficiency		**1.0000**	**1.0189**	**1.0746**	**1.1087**	**1.3802**
β_3	Mean/Bias	0.1	0.1213/0.0213	0.1161/0.0161	0.1159/0.0159	0.1225/0.0225	0.1130/0.0130
	SD		0.3080	0.3144	0.3053	0.3006	0.2713
	MSE		0.0953	0.0991	0.0935	0.0909	0.0738
	Relative Efficiency		**1.0000**	**0.9596**	**1.0178**	**1.0497**	**1.2887**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2645/0.0145	0.2805/0.0305	0.2859/0.0359	0.2957/0.0457
	SD		0.4451	0.4421	0.4264	0.3999	0.3377
	MSE		0.1982	0.1957	0.1828	0.1612	0.1162
	Relative Efficiency		**1.0000**	**1.0137**	**1.0897**	**1.2390**	**1.7370**
γ_1	Mean/Bias	-0.1	-0.0937/0.0063	-0.0854/0.0146	-0.1076/-0.0076	-0.0972/0.0028	-0.1019/-0.0019
	SD		0.5371	0.5384	0.5166	0.4705	0.3996
	MSE		0.2885	0.2901	0.2669	0.2214	0.1597
	Relative Efficiency		**1.0000**	**0.9951**	**1.0809**	**1.3027**	**1.8062**
γ_2	Mean/Bias	0.5	0.4992/-0.0008	0.5142/0.0142	0.5281/0.0281	0.5370/0.0370	0.5181/0.0181
	SD		0.5823	0.5886	0.5648	0.5149	0.4345
	MSE		0.3391	0.3467	0.3198	0.2665	0.1891
	Relative Efficiency		**1.0000**	**0.9785**	**1.0628**	**1.2789**	**1.7962**
γ_3	Mean/Bias	-0.1	-0.0982/0.0018	-0.0956/0.0044	-0.1056/-0.0056	-0.1178/-0.0178	-0.0981/0.0019
	SD		0.4735	0.4784	0.4522	0.4318	0.3401
	MSE		0.2242	0.2289	0.2045	0.1868	0.1157
	Relative Efficiency		**1.0000**	**0.9795**	**1.0964**	**1.2025**	**1.9380**

Figure A.16: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity

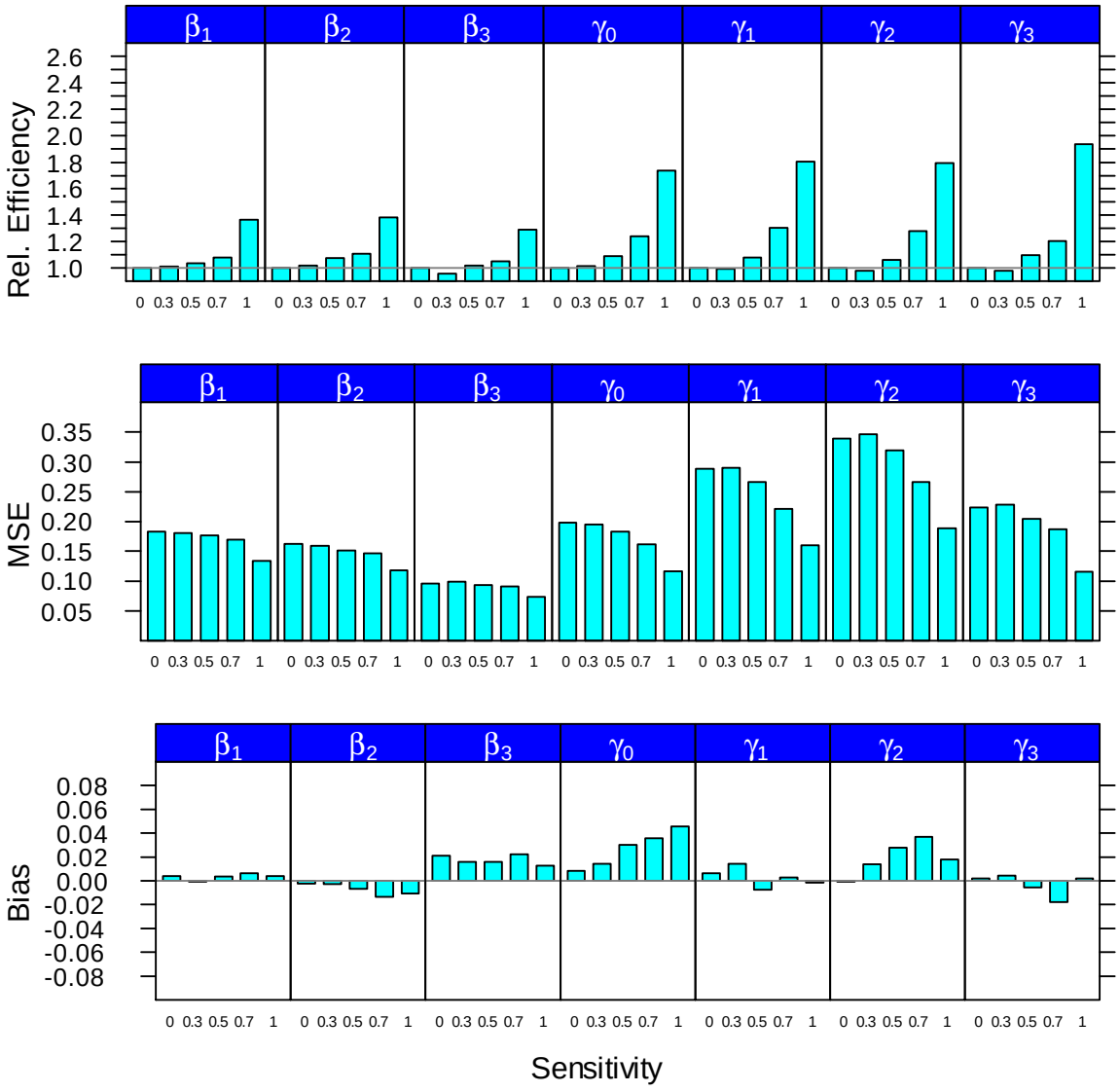


Table A.17: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2037/0.0037	0.2027/0.0027	0.2053/0.0053	0.2024/0.0024	0.1981/-0.0019
	SD		0.4272	0.4273	0.4201	0.4131	0.3887
	MSE		0.1825	0.1826	0.1765	0.1706	0.1511
	Relative Efficiency		**1.0000**	**0.9996**	**1.0342**	**1.0695**	**1.2076**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3039/-0.0039	-0.3031/-0.0031	-0.3038/-0.0038	-0.3082/-0.0082
	SD		0.4029	0.3977	0.3930	0.3817	0.3580
	MSE		0.1623	0.1582	0.1545	0.1457	0.1283
	Relative Efficiency		**1.0000**	**1.0261**	**1.0509**	**1.1141**	**1.2664**
β_3	Mean/Bias	0.1	0.1213/0.0213	0.1204/0.0204	0.1215/0.0215	0.1191/0.0191	0.1215/0.0215
	SD		0.3080	0.3073	0.3053	0.2990	0.2836
	MSE		0.0953	0.0948	0.0937	0.0898	0.0809
	Relative Efficiency		**1.0000**	**1.0046**	**1.0177**	**1.0609**	**1.1791**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2515/0.0015	0.2514/0.0014	0.2495/-0.0005	0.2401/-0.0099
	SD		0.4451	0.4348	0.4106	0.3981	0.3516
	MSE		0.1982	0.1890	0.1686	0.1584	0.1237
	Relative Efficiency		**1.0000**	**1.0481**	**1.1752**	**1.2505**	**1.6025**
γ_1	Mean/Bias	-0.1	-0.0937/0.0063	-0.0871/0.0129	-0.0920/0.0080	-0.0902/0.0098	-0.0861/0.0139
	SD		0.5371	0.5251	0.5012	0.4761	0.4222
	MSE		0.2885	0.2759	0.2513	0.2268	0.1785
	Relative Efficiency		**1.0000**	**1.0460**	**1.1483**	**1.2725**	**1.6181**
γ_2	Mean/Bias	0.5	0.4992/-0.0008	0.5039/0.0039	0.4985/-0.0015	0.4941/-0.0059	0.4996/-0.0004
	SD		0.5823	0.5643	0.5425	0.5177	0.4578
	MSE		0.3391	0.3185	0.2943	0.2680	0.2096
	Relative Efficiency		**1.0000**	**1.0647**	**1.1521**	**1.2651**	**1.6175**
γ_3	Mean/Bias	-0.1	-0.0982/0.0018	-0.0929/0.0071	-0.0945/0.0055	-0.0883/0.0117	-0.0879/0.0121
	SD		0.4735	0.4673	0.4406	0.4151	0.3722
	MSE		0.2242	0.2184	0.1942	0.1724	0.1387
	Relative Efficiency		**1.0000**	**1.0266**	**1.1547**	**1.3013**	**1.6183**

Figure A.17: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Known Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information

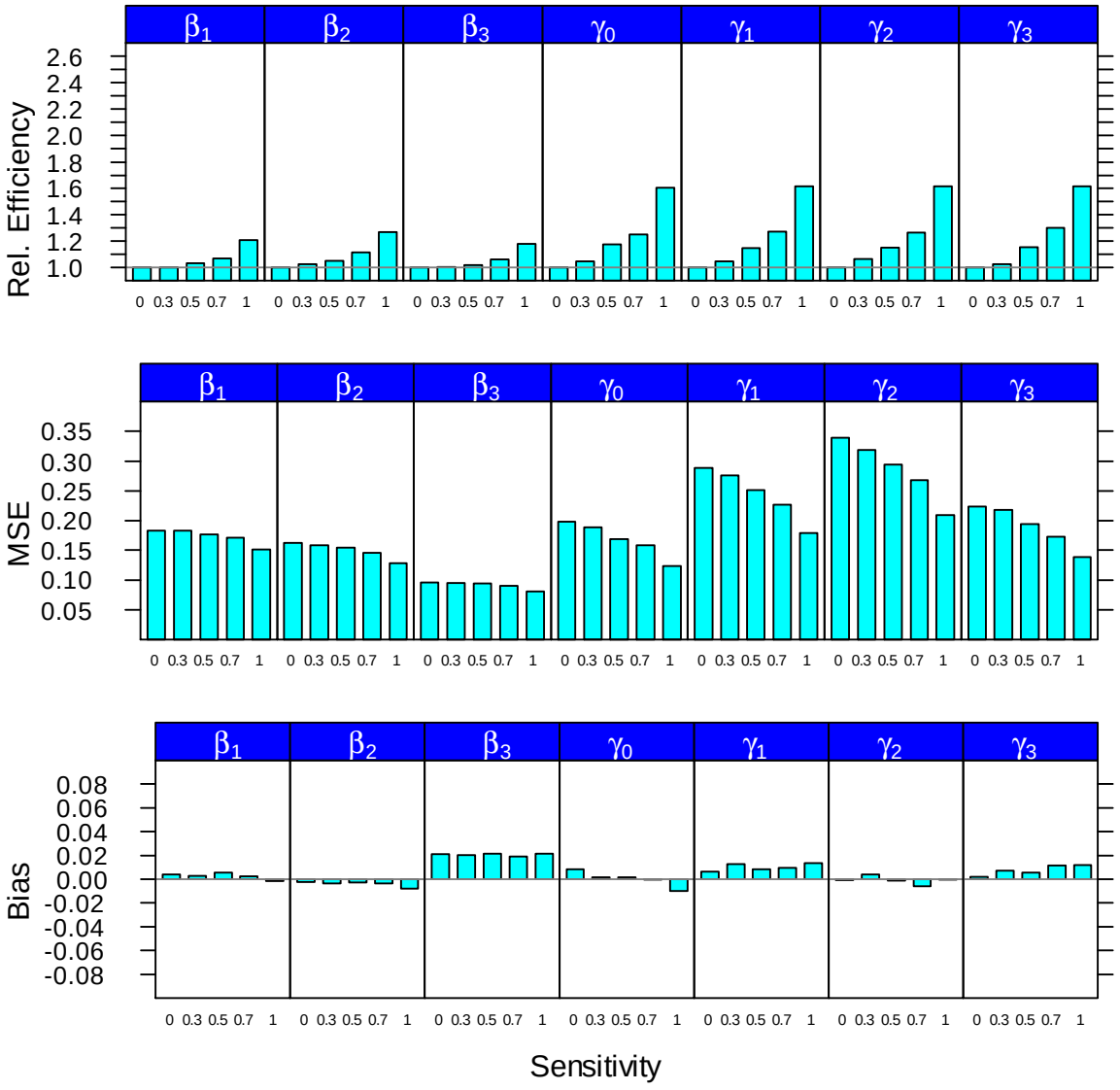
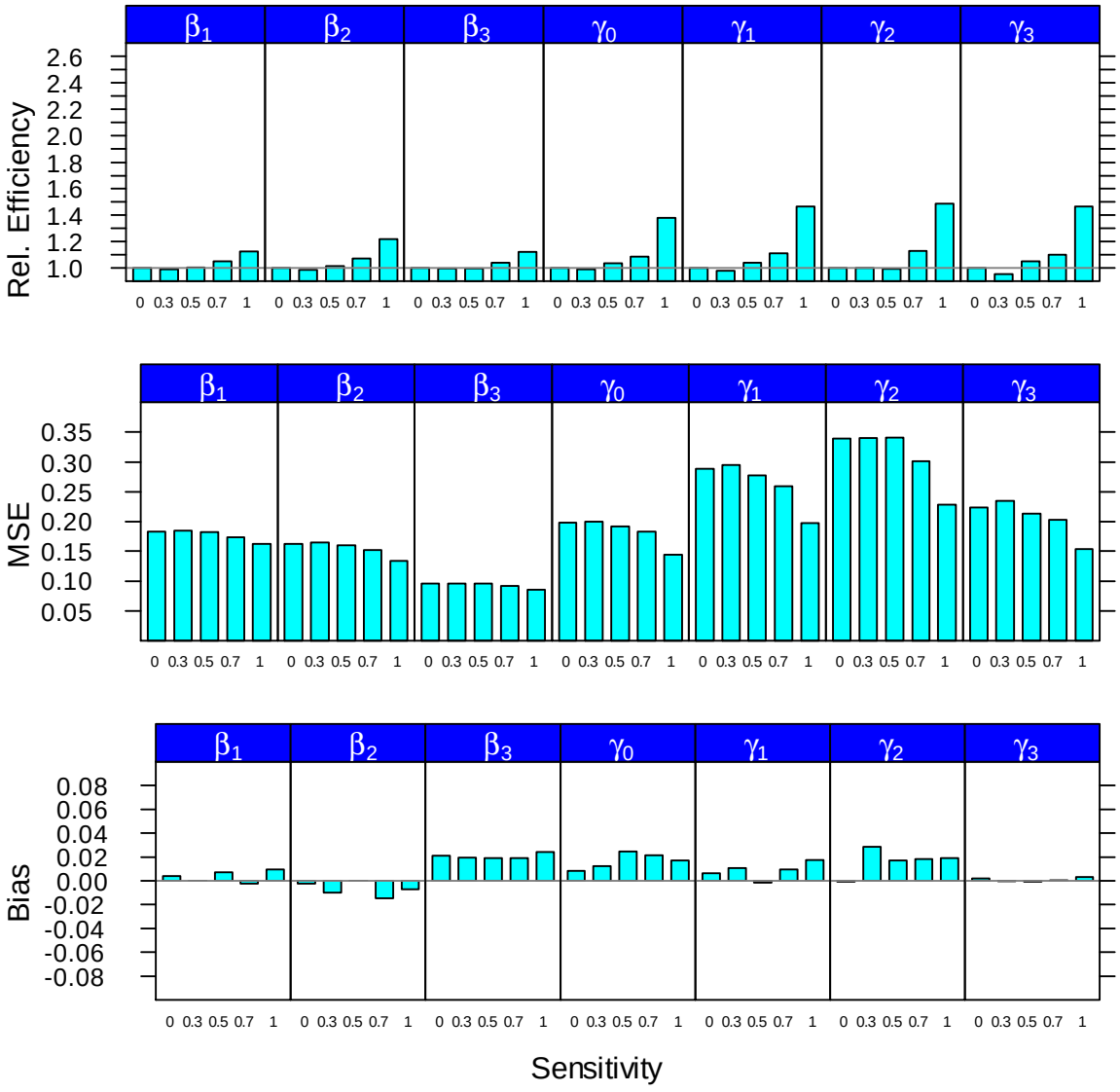


Table A.18: PH Simulation Results for 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information

	Statistics	True Parameter	No Known Cure Info	p_0			
				0.3	0.5	0.7	1
β_1	Mean/Bias	0.2	0.2037/0.0037	0.1997/-0.0003	0.2071/0.0071	0.1975/-0.0025	0.2095/0.0095
	SD		0.4272	0.4291	0.4263	0.4165	0.4028
	MSE		0.1825	0.1841	0.1818	0.1735	0.1624
	Relative Efficiency		**1.0000**	**0.9911**	**1.0042**	**1.0520**	**1.1246**
β_2	Mean/Bias	-0.3	-0.3026/-0.0026	-0.3102/-0.0102	-0.3000/0.0000	-0.3146/-0.0146	-0.3074/-0.0074
	SD		0.4029	0.4053	0.3997	0.3890	0.3651
	MSE		0.1623	0.1644	0.1598	0.1515	0.1334
	Relative Efficiency		**1.0000**	**0.9881**	**1.0160**	**1.0729**	**1.2175**
β_3	Mean/Bias	0.1	0.1213/0.0213	0.1196/0.0196	0.1194/0.0194	0.1193/0.0193	0.1244/0.0244
	SD		0.3080	0.3084	0.3086	0.3022	0.2908
	MSE		0.0953	0.0955	0.0956	0.0917	0.0852
	Relative Efficiency		**1.0000**	**0.9974**	**0.9959**	**1.0388**	**1.1218**
γ_0	Mean/Bias	0.25	0.2582/0.0082	0.2626/0.0126	0.2748/0.0248	0.2715/0.0215	0.2671/0.0171
	SD		0.4451	0.4474	0.4373	0.4273	0.3794
	MSE		0.1982	0.2003	0.1919	0.1830	0.1442
	Relative Efficiency		**1.0000**	**0.9898**	**1.0359**	**1.0852**	**1.3767**
γ_1	Mean/Bias	-0.1	-0.0937/0.0063	-0.0890/0.0110	-0.1019/-0.0019	-0.0902/0.0098	-0.0823/0.0177
	SD		0.5371	0.5427	0.5269	0.5096	0.4441
	MSE		0.2885	0.2947	0.2776	0.2598	0.1976
	Relative Efficiency		**1.0000**	**0.9792**	**1.0390**	**1.1108**	**1.4622**
γ_2	Mean/Bias	0.5	0.4992/-0.0008	0.5286/0.0286	0.5174/0.0174	0.5183/0.0183	0.5194/0.0194
	SD		0.5823	0.5823	0.5837	0.5484	0.4779
	MSE		0.3391	0.3399	0.3411	0.3011	0.2288
	Relative Efficiency		**1.0000**	**1.0000**	**0.9950**	**1.1274**	**1.4844**
γ_3	Mean/Bias	-0.1	-0.0982/0.0018	-0.1007/-0.0007	-0.1008/-0.0008	-0.0997/0.0003	-0.0968/0.0032
	SD		0.4735	0.4842	0.4622	0.4512	0.3913
	MSE		0.2242	0.2345	0.2137	0.2036	0.1531
	Relative Efficiency		**1.0000**	**0.9562**	**1.0493**	**1.1013**	**1.4644**

Figure A.18: Bias and Relative Efficiency for the PH Simulation Results of 90% Specificity ($p_1 = 0.1$), and Baseline Weibull Hazard ($h = \frac{1}{3}$, $k = 4$) – Assume Unknown Sensitivity and Specificity, 50% Subjects with Missing Diagnostic Information



References

Alan B. Cantor and Jonathan J. Shuster, 1992, Parametric versus non-parametric methods for estimating cure rates based on censored survival data, Statistics in Medicine, Volume 11:931 - 937.

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.

Cox, D.R., 1972. Regression models and life-tables. Journal of Royal Statistical Society, Serie B 34, 187–220.

Cox, D.R., 1975. Partial Likelihood. Biometrika, 62, 269-276.

Farewell, V.T., 1982, The use of mixture models for the analysis of survival data with long-term survivors, Biometrics, Vol. 38, No. 4., pp. 1041-1046.

Gamel, J. W., McLean, I. W. and Rosenberg, S. H., 1990, Proportion cured and mean log survival time as functions of tumor size, Statistics in Medicine, 9, 999-1006.

Ghitany, M. E. and Maller, R. A., 1992, Asymptotic results for exponential mixture models with long term survivors, Statistics, 23, 321-336.

Goldman, A. I., 1984, Survivorship analysis when cure is a possibility: A Monte Carlo study, Statistics in Medicine, 3, 153-163.

Gordon, N. H., 1990, Maximum likelihood estimation for mixtures of two gompertz distributions when censoring occurs, Communications in Statistics - Simulation and Computation, 19, 733-747.

Gordon, N. H. 1990, Application of the theory of finite mixtures for the estimation of cure rates of treated cancer patients, Statistics in Medicine, 9:397-407.

Gupta N, Wang CC, Jones CU, Doppke KP, Myint S, Thompson M, Parsons JT, Mendenhall WM, Dische S, Aird EGA, Henk JM, Bidmead MAM, Svoboda V, Chan Y, Hanlon AL, Peters TL, Hanks GE. Local control of carcinoma of the tonsil by radiation therapy: an analysis of patterns of fractionation in nine institutions. International Journal of Radiation Oncology, Biology, Physics 1995; 33:549–562.

Hogg, S. A. and Ciampi, A., GFREG: A computer program for maximum likelihood regression using the generalized F distribution, Computer Methods and Programs in Biomedicine, 20, 201-215 (1985).

Horn, R. A.and Johnson, C. R. (1985), Matrix Analysis, Cambridge University Press, ISBN 978-0-521-38632-6.

Jones, D. R., Powles, R. L., Machin, D. and 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.

Leary, J. T., Handling, M., Talerico, M., Lin, Y., Bowe, J. A., 2009, Physeal fractures of the distal tibia: predictive factors of premature physeal closure and growth arrest. Journal of Pediatric Orthopaedics 29 Issue 4, 356-361.

Li CS, Taylor JMG, 2002. A semi-parametric accelerated failure time cure model. Statistics in Medicine; 21:3235–3247.

Louis, T.A., 1982. Finding the observed information matrix when using the EM algorithm. Journal of the Royal Statistical Society, Series B 44 (2), 226–233.

Kalbfleisch, J. D. and Prentice, R. L., The Statistical Analysis of Failure Time Data, Wiley, New York,1980.

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.

Peng, Y., Dear, K.B.G., 2000. A nonparametric mixture model for cure rate estimation. Biometrics 56; 237–243.

Peng Y., 2003a. Fitting semiparametric cure models. Computational Statistics & Data Analysis; 481-490.

Peng Y., 2003b. Estimating baseline distribution in proportional hazards cure models. Computational Statistics & Data Analysis; 187-201.

Ritov Y. Estimation in a linear regression model with censored data. The Annals of Statistics 1990; 18(1):303–328.

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, Vol. 51, No. 3., pp. 899-907.

Wei LJ. The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis (Disc: P1881–1885). Statistics in Medicine 1992; 11:1871–1879.

Withers HR, Peters LJ, Taylor JMG, Owen JB, Morrison WH, Schultheiss TE, Keane T, O'Sullivan B, van Dyk J, Fang HB, Li G, Sun J. Maximum likelihood estimation in a semiparametric logistic/proportional-hazards mixture model. Scandinavian Journal of Statistics. Theory and Applications 2005; 32:59–75.

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 (418), 284–292.

Zhang J., Peng Y, 2007. A new estimation method for the semiparametric accelerated failure time mixture cure model. Statistics in Medicine; 26:3157–3171.

Vita

Yu Wu

EDUCATION

2005-2010 Ph. D. in Biostatistics, University of Medicine & Dentistry of New Jersey

1997-1999 M.S. in Statistics from University of Connecticut, Storrs, CT.

1992-1997 B.S. in Electrical Engineering from Tsinghua University, Beijing, P.R. China.

EXPERIENCE

2006-2010 Statistical Consultant, Akros Pharma, Princeton, NJ.

2005-2006 Statistical Consultant, Mitsubishi Pharma, Warren, NJ.

2004-2005 Programming Consultant, Pfizer, Morris Plains, NJ.

2003-2004 Programming Consultant, Forest Laboratories, Jersey City, NJ.

2002-2003 Database Consultant, Hoffman La Roche, Nutley, NJ.

2000-2002 Senior Programmer II, Forest Laboratories, New York, NY.

1999-2000 Statistical Programmer, Kendle, Cranford, NJ.

1997-1999 Teaching assistant, Department of Statistics, University of Connecticut, Storrs, CT.